JEAN-FRANÇOIS QUESSY

MÉTHODOLOGIE ET APPLICATION DES COPULES: TESTS D'ADÉQUATION, TESTS D'INDÉPENDANCE, ET BORNES POUR LA VALEUR–À-RISQUE

Thèse présentée
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de doctorat en mathématiques
pour l'obtention du grade de Philosophiæ Doctor (Ph.D.)

FACULTÉ DES SCIENCES ET DE GÉNIE
UNIVERSITÉ LAVAL
QUÉBEC

MARS 2005

RÉSUMÉ

Dans la théorie moderne de la dépendance stochastique, les copules s'avèrent un outil de choix. En effet, une copule contient toute l'information pertinente au sujet de la structure de dépendance d'un vecteur de variables aléatoires continues et permet ainsi de l'isoler des effets des lois marginales. Un aspect important de l'application de la théorie des copules concerne le choix d'une famille de modèles qui s'ajustent adéquatement à des observations multivariées. Dans le cas bivarié, ce problème a été abordé par Shih (1998) et Glidden (1999) pour le modèle de fragilité gamma, alors que Genest & Rivest (1993) ont proposé une méthode de sélection graphique applicable à une classe plus large de copules, dites archimédiennes.

Dans la première partie de la thèse, des statistiques d'adéquation applicables à de nombreux modèles de copules à $d \geq 2$ variables sont développés. Les résultats de Barbe et al. (1996) sur la convergence faible du processus de Kendall permettent de caractériser la limite sous l'hypothèse nulle d'un processus empirique proposé pour l'adéquation. Ceci justifie la définition de statistiques de type Cramér–von Mises et Kolmogorov–Smirnov pour l'adéquation dont les seuils s'obtiennent par *bootstrap paramétrique*. La conception de tests basés sur d'autres processus, en particulier le processus de copule empirique proposé et étudié par Deheuvels (1981a,b,c), ainsi que par Gänssler & Stute (1987), est aussi brièvement abordée.

Dans le deuxième segment de la thèse, la performance asymptotique de tests pour l'indépendance multivariée est traitée. Le comportement dans un voisinage de l'indépendance du processus de copule empirique permet de calculer la fonction de puissance et l'efficacité relative asymptotique locale de plusieurs statistiques de tests qui en découlent. Une attention spéciale est portée à quelques statistiques de Cramér–von Mises proposées par Deheuvels (1981a,b,c). En particulier, une statistique pour l'indépendance bivariée inspirée de Blum et coll. (1961) est comparée à la statistique linéaire de rangs localement la plus puissante pour une contre-hypothèse de dépendance caractérisée par un modèle de copules donné. L'efficacité asymptotique locale de statistiques pour l'indépendance multivariée obtenues d'une décomposition de Möbius du processus de copule empirique est également étudiée.

Enfin, la thèse se conclut par l'obtention de bornes sur la valeur-à-risque (VaR) de la somme de risques dépendants en présence d'information partielle. Des bornes explicites sont obtenues lorsque les marges sont connues et quand seuls les deux premiers moments des risques sont disponibles.

REMERCIEMENTS

Je tiens à souligner l'influence importante de plusieurs personnes dans l'accomplissement de ce doctorat. D'abord, un merci du fond du cœur à mes parents, Claude et Solange, qui m'ont toujours encouragé à aller jusqu'au bout et m'ont appuyé inconditionnellement durant ces longues années d'études. J'aimerais aussi exprimer une grande reconnaissance envers une personne toute spéciale pour moi, ma douce moitié Marie-Ève. Par sa patience et sa compréhension à l'endroit d'un conjoint souvent plongé dans ses équations, par ses bons mots d'encouragement, et surtout par son amour, elle a su à tout moment remettre les choses dans leur juste perspective, et ainsi contribuer à ce que ce travail se réalise dans le bonheur.

Aussi, je ne peux et ne pourrai, tout au long de ma carrière, faire abstraction de l'influence énorme et positive qu'ont exercée sur moi Bruno Rémillard et Christian Genest. En acceptant de diriger mes travaux de recherche à la maîtrise et au doctorat, ils auront influencé de façon significative mon parcours autant professionnel que personnel. Leur joie communicative à faire de la recherche, leur curiosité intellectuelle, leur souci de rigueur et leur dynamisme, qui ne se démentent jamais, marqueront à jamais le déroulement de ma carrière. Bruno, tu as cru, bien avant que cette idée ne m'effleure l'esprit, que je pourrais devenir professeur. Je te dois donc une grande partie du bonheur que j'éprouve à exercer ce métier stimulant. Christian, tes mots d'encouragement dans les moments plus difficiles m'ont dévoilé un être humain de grande qualité derrière le chercheur. Je voudrais également remercier mon ami et collègue à l'Université du Québec à Trois-Rivières, Mhamed Mesfioui, qui a facilité mon intégration dans mes nouvelles fonctions et qui trouve toujours le bon mot pour m'encourager.

Je veux de plus remercier le *Conseil de recherche en sciences naturelles et en génie du Canada* (CRSNG) pour la bourse de doctorat qu'il m'a accordée. De même, une partie de ce travail a été financée par des octrois individuels et collectifs consentis à MM. Christian Genest et Bruno Rémillard par le CRSNG, ainsi que par le *Fonds québécois de la recherche sur la nature et les technologies* (FQRNT).

J'exprime finalement ma gratitude envers les membres du jury, MM. Christian Genest, Richard Lockhart, Jean-Claude Massé, Bruno Rémillard et Radu Theodorescu, qui ont accepté de lire et d'évaluer cette thèse. Un remerciement spécial va à M. Theodorescu pour avoir consenti à en effectuer une prélecture.

AVANT-PROPOS

La majeure partie de cette thèse est constituée de cinq articles, écrits en anglais, dont quatre sont le résultat de collaborations avec mon directeur de recherche Christian Genest, de l'Université Laval, et mon co-directeur de recherche Bruno Rémillard, affilié à l'École des HEC de Montréal. Un de ces articles a été accepté pour publication dans le *Journal of Multivariate Analysis*, alors que les autres sont des travaux soumis à différentes revues spécialisées en statistique. Un cinquième papier, préparé conjointement avec M. Mhamed Mesfioui, de l'Université du Québec à Trois-Rivières, paraîtra dans la revue *Insurance: Mathematics and Economics*.

La thèse est composée de sept chapitres, incluant une introduction au chapitre 1. Elle peut se subdiviser en trois blocs relativement homogènes: la partie I traite de tests d'adéquation pour des modèles de copules; la partie II aborde l'efficacité asymptotique locale de tests de Cramér–von Mises pour l'indépendance multivariée; la partie III s'attarde à obtenir des bornes sur la valeur-à-risque (VaR) de la somme de risques dépendants. Le tableau de la page suivante permet d'avoir une vision claire de la structure du document.

Les résumés en français au début des chapitres contenant un article, c'est-à-dire les chapitres 2, 3, 5, 6 et 7, permettent d'assurer une certaine cohérence entre les différents constituants de la thèse. Finalement, pour alléger le texte, toutes les annexes rattachées originellement aux articles ont été confinées à la fin du présent document, dans les annexes A à D.

Table 1: Structure de la thèse et état actuel des articles

Partie	**Chapitre**	**Sujet**	**État**
I	2	*Tests d'adéquation basés sur le processus de Kendall*	ARTICLE SOUMIS
	3	*Correction de la variance asymptotique de la statistique d'adéquation de Shih (1998)*	ARTICLE SOUMIS
	4	*Tests d'adéquation basés sur le processus de copule empirique*	TRAVAIL INÉDIT
II	5	*Comportement asymptotique local d'une statistique de Cramér–von Mises pour l'indépendance bivariée*	ARTICLE ACCEPTÉ
	6	*Comportement asymptotique local de statistiques de Cramér–von Mises pour l'indépendance multivariée*	ARTICLE SOUMIS
III	7	*Bornes explicites sur la valeur-à-risque (VaR) pour la somme de risques dépendants*	ARTICLE ACCEPTÉ

Contents

LISTE DES TABLEAUX

LISTE DES FIGURES

CHAPITRE 1

INTRODUCTION

Pour étudier le comportement simultané des composantes d'un vecteur aléatoire $X = (X_1, \ldots, X_d)$, on peut faire appel à la fonction de répartition jointe

$$H(x_1, \ldots, x_d) = P\left(X_1 \leq x_1, \ldots, X_d \leq x_d\right),$$

qui caractérise totalement le comportement stochastique de X. Par exemple, dans plusieurs applications de l'analyse multivariée classique, on émet souvent l'hypothèse que H est la fonction de répartition d'une loi normale multivariée. Cet usage du modèle normal est souvent justifié par la simplicité relative des calculs à effectuer. Toutefois il est notoire, notamment en économie, que la loi normale s'ajuste très mal à certains types de données.

Dans l'étude moderne de la dépendance, on utilise fréquemment le fait que toute l'information à propos de la structure de dépendance de X est contenue dans une fonction appelée la copule. Le concept de copule est un outil puissant et flexible puisqu'il permet de modéliser la dépendance sans tenir compte de l'effet du comportement des marges, c'est-à-dire des fonctions de répartition $F_1, \ldots, F_d$ des variables $X_1, \ldots, X_d$ prises individuellement.

Formellement, une copule à d dimensions est une fonction de répartition définie sur l'hypercube $[0,1]^d$ et dont les marges sont uniformes.

Le résultat suivant, dû à Sklar (1959), permet de relier la notion de copule à celle de fonction de répartition multivariée.

Théorème 1.1. *Si H est une fonction de répartition d-variée de marges $F_1, \ldots, F_d$, alors il existe une copule $C : [0,1]^d \to [0,1]$ telle que*

$$H(x_1, \ldots, x_d) = C\left\{F_1(x_1), \ldots, F_d(x_d)\right\}. \tag{1.1}$$

Si les distributions marginales $F_1, \ldots, F_d$ sont continues, alors C est unique.

De ce fameux résultat, on déduit l'unique copule associée à une distribution H de marges continues en effectuant le changement de variables $u_i = F_i(x_i)$, $1 \leq i \leq d$. On tire ainsi de la formule (1.1) que

$$C(u_1, \ldots, u_d) = H\left\{F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\right\}. \tag{1.2}$$

À titre d'illustration, considérons la fonction de répartition associée à une loi normale multivariée de moyennes nulles, de variances unitaires et de matrice de corrélation Σ définie positive. Dans ce cas, si N dénote la fonction de répartition d'une variable $\mathcal{N}(0,1)$, alors toutes les marges sont déterminées par N. On en déduit alors, de l'équation (1.2), que la forme de la copule gaussienne est

$$C_\Sigma(u_1, \ldots, u_d) = \int_{-\infty}^{N^{-1}(u_1)} \cdots \int_{-\infty}^{N^{-1}(u_d)} \frac{|\Sigma|^{-d/2}}{(2\pi)^{d/2}} \exp\left(-\frac{1}{2} s^\top \Sigma^{-1} s\right) \mathrm{d}s. \tag{1.3}$$

Une autre classe intéressante de modèles est la famille des copules archimédiennes, dont les membres peuvent s'écrire sous la forme

$$C(u_1, \ldots, u_d) = \phi^{-1}\left\{\phi(u_1) + \cdots + \phi(u_d)\right\},$$

où $\phi : [0,1] \to [0,\infty)$ est un générateur tel que $\phi(1) = 0$ et

$$(-1)^i \frac{\mathrm{d}^i}{\mathrm{d}x^i}\, \phi^{-1}(x) > 0$$

pour tout $i \in \{1, \ldots, d\}$.

Par exemple, si on définit, pour chaque $\theta \geq 0$, le générateur

$$\phi_\theta(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right),$$

on retrouve le modèle paramétrique multivarié de Clayton, c'est-à-dire

$$C_\theta(u_1, \ldots, u_d) = \left(u_1^{-\theta} + \cdots + u_d^{-\theta} - d + 1\right)^{-1/\theta}.$$

Un bénéfice majeur tiré de l'utilisation des copules tient à la possibilité, partant d'un modèle C donné, de bâtir des lois multivariées avec les marges désirées $F_1, \ldots, F_d$. On acquiert ainsi une grande flexibilité quant à l'élaboration de modèles. Une structure de dépendance gaussienne avec des marges de loi Student à différents degrés de liberté, par exemple, pourrait être envisagée par un praticien pour ajuster à des données multivariées.

Une autre conséquence du théorème de Sklar est que toute mesure de dépendance appropriée devrait s'exprimer uniquement en fonction de la copule, car aucune information sur la dépendance n'est incluse dans les marges. Cette propriété est partagée, entre autres, par le rho de Spearman et le tau de Kendall bivariés, définis respectivement par

$$\rho_S = 12 \int_0^1 \int_0^1 \{C(u_1, u_2) - u_2 u_1\} \, \mathrm{d}u_2 \, \mathrm{d}u_1$$

et

$$\tau = 4 \int_0^1 \int_0^1 C(u_1, u_2) \, \mathrm{d}C(u_1, u_2) - 1.$$

Tests d'adéquation pour des modèles de copules

Considérons un échantillon aléatoire multivarié de taille n, à savoir

$$(X_{11}, \ldots, X_{1d}), \ldots, (X_{n1}, \ldots, X_{nd}),$$

et définissons les vecteurs de rangs associés à chaque observation par

$$(R_{11}, \ldots, R_{1d}), \ldots, (R_{n1}, \ldots, R_{nd}),$$

où R_{ij} est le rang de X_{ij} parmi $X_{1j}, \ldots, X_{nj}$. Pour modéliser la structure de dépendance sous-jacente à ces observations, on émet généralement l'hypothèse que ces données proviennent d'une famille paramétrique de copules $\mathcal{C} = (C_\theta)$ indicée par un paramètre $\theta \in \mathcal{O} \subseteq \mathbb{R}^m$. Ensuite, un membre approprié de cette classe est choisi. En général, cela revient à considérer un estimateur convergent θ_n pour le paramètre inconnu θ, et à sélectionner le modèle C_{θ_n}. Cette approche est dite semi-paramétrique, car aucune hypothèse n'est émise sur les marges, quoique la famille de modèles possibles soit imposée.

Les conclusions émises suite à une modélisation obtenue par l'approche semi-paramétrique décrite précédemment doivent cependant être considérées avec précaution. En effet, des conclusions erronées peuvent surgir dans le cas où la vraie copule d'une population n'appartient pas à la famille qui a été supposée au départ. Un premier objectif de cette thèse est de développer des tests d'adéquation dont le but consiste à vérifier si la copule sous-jacente à une population appartient à une certaine famille paramétrique $\mathcal{C} = (C_\theta)$. En d'autres termes, les hypothèses

$$H_0 : C \in \mathcal{C} \quad \text{et} \quad H_1 : C \notin \mathcal{C}$$

seront confrontées. Toutes les statistiques de tests proposées dans cette thèse sont des fonctionnelles de processus empiriques bâtis à partir des rangs des observations.

Au chapitre 2, un premier test d'adéquation à une famille $\mathcal{C} = (C_\theta)$ est suggéré. Celui-ci est basé sur la transformation intégrale de probabilité définie, pour chaque membre C_θ de la famille $\mathcal{C}$, par

$$K_\theta(t) = \mathrm{P}\left\{C_\theta\left(U_1, \ldots, U_d\right) \leq t\right\},$$

où $(U_1, \ldots, U_d)$ est un vecteur aléatoire de loi C_θ. En général, la transformation intégrale de probabilité de C_θ ne caractérise pas complètement les membres d'une classe de copules puisque K_θ constitue une projection dans un espace de fonctions à une variable d'une fonction à d composantes. Néanmoins, il existe quelques exceptions. Par exemple, pour une copule archimédienne bivariée de générateur ϕ, on peut montrer que

$$K(t) = t - \frac{\phi(t)}{\phi'(t)},$$

d'où il s'ensuit que

$$\phi(t) = \exp\left\{\int_{t_0}^{t} \frac{1}{s - K(s)}\,\mathrm{d}s\right\},$$

où $t_0 \in \mathbb{R}$ est arbitraire. On constate donc que la transformation intégrale de probabilité définie entièrement les copules archimédiennes bivariées.

Genest & Rivest (1993) ont proposé comme estimateur de K_θ la fonction non paramétrique

$$K_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\left(e_{i,n} \le t\right), \tag{1.4}$$

où

$$e_{i,n} = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}\left(R_{j1} \le R_{i1}, \ldots, R_{jd} \le R_{id}\right).$$

Sous certaines conditions vérifiées pour un grand nombre de familles de copules fréquemment utilisées en pratique, Barbe et coll. (1996) ont obtenu la convergence du processus empirique

$$\mathbb{K}_{n,\theta}(t) = \sqrt{n}\left\{K_n(t) - K_\theta(t)\right\}$$

vers une limite gaussienne centrée $\mathbb{K}_\theta$, lorsque l'échantillon qui sert à estimer K_θ est tiré d'une population dont la copule est C_θ. Afin de tester l'adéquation à une famille de copules, ce résultat sera étendu ici au processus d'adéquation

$$\mathbb{K}_n = \sqrt{n}\left\{K_n(t) - K_{\theta_n}(t)\right\},$$

dans lequel θ_n est un estimateur convergent de θ. Au prix d'une hypothèse supplémentaire imposée aux membres de $\mathcal{C}$, il sera montré que $\mathbb{K}_n$ converge vers une limite gaussienne centrée de représentation

$$\mathbb{K}(t) = \mathbb{K}_\theta(t) - \Theta \dot{K}_\theta(t),$$

où Θ est la limite en loi de $\Theta_n = \sqrt{n}(\theta_n - \theta)$ et $\dot{K}_\theta$ est la dérivée partielle de K_θ par rapport à θ. Ce résultat asymptotique permet d'identifier la loi limite des statistiques d'adéquation de type Cramér–von Mises et Kolmogorov–Smirnov définies par

$$S_n = \int_0^1 \mathbb{K}_n^2(t)\,\mathrm{d}K_{\theta_n}(t) \quad \text{et} \quad T_n = \sup_{0\le t\le 1} |\mathbb{K}_n(t)|\,.$$

En effet, on déduit de la convergence de $\mathbb{K}_n$ vers $\mathbb{K}$ que

$$S_n \rightsquigarrow S = \int_0^1 \mathbb{K}^2(t)\,\mathrm{d}K_{\theta}(t) \quad \text{et} \quad T_n \rightsquigarrow T = \sup_{0\le t\le 1} |\mathbb{K}(t)|\,,$$

où $\rightsquigarrow$ dénote la convergence faible, aussi dite convergence en loi. Toutefois, comme les distributions de S et T sous H_0 ne sont pas explicites et qu'en plus, elles dépendent de la valeur inconnue de θ, la méthode d'auto-échantillonnage ou de *bootstrap* paramétrique sera employée pour calculer des seuils asymptotiquement exacts pour les tests basés sur S_n et T_n. Il s'agit en fait de générer N échantillons de taille n tirés de la loi C_{θ_n}, et de calculer à chaque fois la statistique d'intérêt. Pour la statistique de Cramér–von Mises, par exemple, le seuil observé estimé correspondant à une valeur observée S_n est donné par

$$\frac{1}{N}\sum_{i=1}^{n} \mathbf{1}\left(S_{ni} > S_n\right),$$

où $S_{n1},\ldots,S_{nN}$ sont les statistiques calculées pour les échantillons bootstrap. Cette méthode est utilisée pour montrer, par simulation, la bonne performance de S_n et T_n du point de vue de la puissance.

Au chapitre 3, on corrige une erreur dans le calcul de la variance asymptotique d'une statistique d'adéquation proposée par Shih (1998). Cette erreur a été décelée lors de l'étude de simulation conduite au chapitre 2. La procédure d'adéquation de Shih, qui s'applique uniquement au modèle bivarié de Clayton, consiste à comparer un estimateur de concordance non pondéré $\tilde{\theta}_n$ proposé par Oakes (1982), à un estimateur pondéré $\hat{\theta}_n$ introduit par Clayton &

Cuzick (1985) et dont le comportement asymptotique a été par étudié par Oakes (1986). Sous l'hypothèse nulle d'appartenance à la famille de Clayton, la statistique de test suggérée par Shih, c'est-à-dire $R_n = \sqrt{n}\log(\tilde{\theta}_n/\hat{\theta}_n)$, est asymptotiquement normale de moyenne nulle et de variance

$$\sigma^2(\theta) = \lim_{n\to\infty} \mathrm{var}(R_n).$$

Toutefois, l'expression donnée par Shih (1998) est erronée puisque $\sigma^2(\theta) < 0$ pour certaines valeurs de θ. Une expression correcte pour $\sigma^2(\theta)$ est obtenue, en utilisant une approche de calcul différente de celle employée par Oakes (1982, 1986) et par Shih (1998).

En général, puisque l'expression analytique des membres de $\mathcal{C}$ ne peut pas être retrouvée en connaissant uniquement K_θ, il est clair que les tests basés sur $\mathbb{K}_n$ ne sont pas toujours convergents. De plus, du fait que K_θ est une projection, on s'attend à ce que l'efficacité des tests basés sur $\mathbb{K}_n$ diminue à mesure que d augmente, sans compter que K_θ ne possède pas d'expression explicite pour certaines familles de copules. Ces constatations motivent à proposer, au chapitre 4, des tests potentiellement plus puissants basés sur l'estimation de C_θ par la copule empirique

$$C_n(u) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\left(\frac{R_{i1}}{n} \le u_1, \ldots, \frac{R_{id}}{n} \le u_d\right),$$

où $u = (u_1, \ldots, u_d)$. Gänssler & Stute (1987) ont obtenu le résultat selon lequel le processus de copule empirique

$$\mathbb{C}_{n,\theta}(u) = \sqrt{n}\left\{C_n(u) - C_\theta(u)\right\},$$

fondé uniquement sur les rangs des observations, converge vers une limite gaussienne centrée $\mathbb{C}_\theta$. En émettant quelques hypothèses supplémentaires sur les membres de la famille $\mathcal{C}$, on obtient que le processus d'adéquation

$$\mathbb{C}_n(u) = \sqrt{n}\,\{C_n(u) - C_{\theta_n}(u)\}$$

converge vers une limite gaussienne centrée de représentation

$$\mathbb{C}(u) = \mathbb{C}(u) - \Theta \dot{C}_\theta(u).$$

Ce résultat assure que les fonctionnelles continues calculées à partir de $\mathbb{C}_n$ seront convergentes, et permet de proposer un test d'adéquation pour la copule normale.

Efficacité asymptotique locale de tests de Cramér–von Mises pour l'indépendance multivariée

Le problème de tester que les composantes d'un vecteur aléatoire multivarié sont indépendantes est courant. De fait, de tels tests devraient être appliqués de manière routinière avant même d'envisager une modélisation de la dépendance, que ce soit à l'aide de copules ou non. Dans le cas bivarié, l'approche classique consiste à baser un test sur le coefficient de corrélation empirique de Pearson, qui peut être défini par

$$r_n = \frac{1}{s_{X_1} s_{X_2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \{H_n(x_1, x_2) - H_n(x_1, \infty) H_n(\infty, x_2)\}\, \mathrm{d}x_2\, \mathrm{d}x_1,$$

où H_n est la fonction de répartition empirique basée sur un échantillon bivarié $(X_{11}, X_{12}), \ldots, (X_{n1}, X_{n2})$, et s_{X_1}, s_{X_2} sont les écart-types expérimentaux. Toutefois, en général, la loi de r_n dépend des marges, ce qui la rend impropre à mesurer adéquatement la dépendance, en dehors du cadre gaussien classique.

L'approche non paramétrique consiste à vérifier que la structure sous-jacente à une population multivariée est caractérisée par la fonction de dépendance

$C_{\theta_0}(u_1, \ldots, u_d) = u_1 \cdots u_d$, c'est-à-dire par la copule d'indépendance. Dans le cas $d = 2$, une classe de statistiques de tests pour vérifier que $C = C_{\theta_0}$ est donnée par la famille des statistiques linéaires de rangs, dont la forme générale est

$$S_n^J = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{R_{1i}}{n+1}, \frac{R_{2i}}{n+1}\right),$$

en terme d'une fonction score $J : [0,1]^2 \to \mathbb{R}$. Le rho de Spearman et la statistique de Van der Waerden appartiennent à cette classe. Des conditions sur la fonction J qui assurent que $\sqrt{n}(S_n^J - \theta_0)$ est asymptotiquement normale sous l'hypothèse nulle d'indépendance se retrouvent, entre autres, dans les travaux de Ruymgaart et coll. (1972).

Genest & Verret (2005), exploitant la normalité asymptotique de S_n^J, ont calculé l'efficacité relative asymptotique de Pitman pour plusieurs paires de statistiques linéaires de rangs et pour différents scénarios de contre-hypothèses contiguës à l'indépendance. Spécifiquement, ces auteurs ont déterminé la loi limite de S_n^J pour des contre-hypothèses du type C_{δ_n}, où $\delta_n = \theta_0 + \delta/\sqrt{n}$.

Les tests basés sur les statistiques linéaires de rangs peuvent toutefois être qualifiés de ponctuels, puisqu'ils s'intéressent à une seule caractéristique de la loi, c'est-à-dire à l'espérance $\mathrm{E}\{J(U_1, U_2)\}$, où (U_1, U_2) est de loi C_θ. Ceci peut conduire à des tests dont la puissance est limitée. Il existe en effet des structures de dépendance différentes de C_{θ_0}, mais pour lesquelles on a $\mathrm{E}(S_n^J) = \theta_0$, ce qui conduit à des tests de faible puissance.

Une idée inspirée d'une suggestion de Blum et coll. (1961) et développée par Deheuvels (1979, 1980) consiste à considérer la statistique de Cramér–von

Mises du processus de copule empirique multivarié sous l'indépendance, à savoir

$$B_n = \int_{(0,1)^d} \{C_n(u) - C_{\theta_0}(u)\}^2 \, \mathrm{d}u.$$

Le test d'indépendance basé sur B_n amène asymptotiquement le rejet de toute contre-hypothèse de copule différente de C_{θ_0}.

Au chapitre 5, l'efficacité asymptotique locale de B_n est étudiée pour le cas bivarié. Le comportement de $\mathbb{C}_n$ sous des hypothèses contiguës permet de caractériser la loi de B_n sous C_{δ_n} et de retrouver les résultats de Genest & Verret (2005) concernant S_n^J. Pour une famille de contre-hypothèses donnée, ceci permet de comparer la courbe de puissance locale de B_n à celle de la statistique linéaire de rangs localement la plus puissante pour cette classe de modèles. Une généralisation de l'efficacité relative asymptotique de Pitman est également proposée afin de comparer la performance locale de B_n versus une statistique linéaire de rangs.

Au chapitre 6, l'efficacité asymptotique locale de tests pour l'indépendance multivariée basés sur des fonctionnelles de Cramér–von Mises est étudiée. En particulier, la version multivariée de B_n est considérée, de même que quatre procédures de test issues de la décomposition de Möbius du processus de copule empirique $\mathbb{C}_n$. Spécifiquement, la transformation de Möbius décompose $\mathbb{C}_n$ en $2^d - d - 1$ processus asymptotiquement indépendants, ce qui permet de construire des statistiques basées sur des combinaisons d'éléments indépendants à la limite et dont la loi asymptotique est connue. Une extension d-variée de la caractérisation de la loi de $\mathbb{C}_n$ sous des contre-hypothèses contiguës permet de calculer, sous divers scénarios de dépendance, les courbes de puissance locale de ces statistiques de même qu'une efficacité relative asymptotique pour certaines paires de tests.

Bornes sur la valeur-à-risque (VaR) de la somme de risques dépendants

La valeur-à-risque (VaR) au niveau α d'une variable aléatoire X de fonction de répartition F est définie par

$$\mathrm{VaR}_\alpha(X) = F^{-1}(\alpha),$$

c'est-à-dire par le quantile d'ordre α de F. La VaR mesure ainsi le niveau de risque associé à un phénomène particulier. Dans de nombreuses applications, on s'intéresse toutefois à la VaR d'un portefeuille composé de risques potentiellement dépendants.

Soient donc les variables aléatoires continues et positives $X_1, \ldots, X_n$ et considérons la somme $S = X_1 + \cdots + X_n$. En pratique, l'hypothèse que les risques sont indépendants est souvent émise, ce qui permet d'obtenir une expression pour $\mathrm{VaR}_\alpha(S)$. Cette hypothèse est cependant rarement justifiable.

De façon plus réaliste, on suppose que la structure de dépendance du vecteur aléatoire $(X_1, \ldots, X_n)$ est caractérisée par une copule C. En présence d'information partielle sur C, on peut néanmoins obtenir des bornes entre lesquelles $\mathrm{VaR}_\alpha(S)$ se situe. Un tel résultat se retrouve dans l'article de Embrechts et coll. (2003), où la VaR d'une fonction ψ de n risques est considérée. Dans le cas présent, c'est-à-dire pour $\psi(x_1, \ldots, x_n) = x_1 + \cdots + x_n$, on obtient comme cas particulier que si des copules C_L et C_U existent telles que $C \geq C_L$ et $C^d \leq C_U^d$, où C^d est le dual de C, alors

$$\underline{\mathrm{VaR}}_{C_L}(\alpha) \leq \mathrm{VaR}_\alpha(s) \leq \overline{\mathrm{VaR}}_{C_U}(\alpha).$$

Le premier objectif du chapitre 7 est d'obtenir des expressions explicites pour $\underline{\mathrm{VaR}}_{C_L}$ et $\overline{\mathrm{VaR}}_{C_U}$ lorsque les fonctions de répartition marginales $F_1, \ldots, F_n$ de chacun des risques sont connues. Ceci est accompli en imposant quelques restrictions sur les densités associées à $F_1, \ldots, F_n$.

En deuxième lieu, une méthode simple pour borner $\mathrm{VaR}_\alpha(S)$ est proposée lorsque seuls les deux premiers moments des risques, à savoir les moyennes $\mu_1, \ldots, \mu_n$ et les variances $\sigma_1^2, \ldots, \sigma_n^2$, sont connues. Les expressions pour les bornes ainsi obtenues sont simples et dépendent uniquement de

$$\mu = \mu_1 + \cdots + \mu_n \quad \text{et} \quad \sigma = \sigma_1 + \cdots + \sigma_n.$$

CHAPITRE 2

GOODNESS-OF-FIT PROCEDURES FOR COPULA MODELS BASED ON THE PROBABILITY INTEGRAL TRANSFORMATION

Résumé

Wang & Wells (2000) décrivent une méthode d'adéquation non paramétrique pour des données censurées applicable aux copules archimédiennes bivariées. Leur procédure est basée sur une version arbitrairement tronquée du processus de Kendall introduit par Genest & Rivest (1993) et étudié ultérieurement par Barbe et coll. (1996). Bien que Wang & Wells (2000) déterminent le comportement asymptotique de leur processus tronqué, la procédure de sélection de modèle qu'ils proposent est fondée exclusivement sur la valeur observée de la norme L^2; aucun test d'adéquation formel n'est élaboré. Cet article montre comment calculer les seuils asymptotiques de statistiques de tests d'adéquation de type Cramér–von Mises et Kolmogorov–Smirnov basées sur une version non-tronquée du processus de Kendall. Il y est également démontré que les conditions sous lesquelles la convergence faible est assurée sont vérifiées pour plusieurs modèles archimédiens et non-archimédiens.

Abstract

Wang & Wells (2000) describe a nonparametric approach for checking whether the dependence structure of a random sample of censored bivariate data is appropriately modelled by a given family of Archimedean copulas. Their procedure is based on an arbitrarily truncated version of the Kendall process introduced by Genest & Rivest (1993) and later studied by Barbe et al. (1996). Although Wang & Wells (2000) determine the asymptotic behavior of their truncated process, the model selection method they propose is based exclusively on the observed value of its L^2–norm; no formal goodness-of-fit test is derived. This paper shows how to compute asymptotic P-values for Cramér–von Mises and Kolmogorov–Smirnov goodness-of-fit test statistics based on a non-truncated version of Kendall's process. The conditions under which weak convergence occurs are seen to hold for various Archimedean and non-Archimedean copula models commonly met in practice. Simulations are used to study the empirical behavior of the proposed goodness-of-fit tests in finite samples, and their application is illustrated on two classical sets of multivariate data. Power comparisons are also made with tests of model adequacy specifically developed by Shih (1998) for the gamma frailty family.

2.1 Introduction

Due in part to their connection with frailty models in survival analysis (Oakes 1989, 2001), Archimedean copulas have become quite popular as a tool for describing the dependence between two random variables X and Y with continuous marginal distributions F and G, respectively. Given a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ with joint cumulative distribution function

$$H(x, y) = C\{F(x), G(y)\}, \quad x, y \in \mathbb{R}$$

there is thus considerable interest in testing whether the unique underlying copula C may be expressed in the form

$$C(u, v) = C_{\phi_\theta}(u, v) \equiv \phi_\theta^{-1}\{\phi_\theta(u) + \phi_\theta(v)\},$$

where (ϕ_θ) is a class of Archimedean generators $\phi_\theta : (0, 1] \mapsto [0, \infty)$ indexed by a parameter $\theta \in \mathbb{R}$ and satisfying the following conditions

$$\phi_\theta(1) = 0, \quad (-1)^i \frac{\mathrm{d}^i}{\mathrm{d}t^i}\, \phi_\theta^{-1}(t) > 0, \quad i \in \{1, 2\}.$$

Using the fact that the distribution function K of the probability integral transformation $V = H(X, Y)$ is of the form

$$K(\theta, t) = t - \frac{\phi_\theta(t)}{\phi'_\theta(t)}\,, \quad t \in (0, 1]$$

whenever $C \in \mathcal{C} = (C_{\phi_\theta})$, Genest & Rivest (1993) proposed a graphical procedure for selecting an Archimedean model through a visual comparison of a nonparametric estimate K_n of K with a parametric estimate $K(\theta_n, \cdot)$ obtained under the composite null hypothesis $H_0 : C \in \mathcal{C}$.

In their work, Genest & Rivest (1993) simply defined K_n as an empirical cumulative distribution function allocating a weight of $1/n$ to each pseudo-observation

$$V_i = \frac{1}{n-1} \# \{j : X_j < X_i, Y_j < Y_i\}.$$

As for $K(\theta_n, \cdot)$, it was obtained by finding the value θ_n of θ such that, under H_0, the population value $\tau(\theta) = 4\mathrm{E}(V) - 1$ of Kendall's tau matches its standard empirical version, given by $\tau_n = 4\bar{V} - 1$, where $\bar{V} = (V_1 + \cdots + V_n)/n$.

By identifying the pointwise limit of Kendall's process $\sqrt{n}\,\{K_n(\cdot) - K(\theta, \cdot)\}$, Genest & Rivest (1993) were able to construct confidence bands to help with the identification of a proper family $\mathcal{C}$. The limit of the process as such was later identified for arbitrary d–dimensional copulas by Barbe et al. (1996).

Restricting themselves to the bivariate Archimedean case, but allowing for censorship, Wang & Wells (2000) furthered this work by proposing a goodness-of-fit statistic

$$S_{\xi n} = \int_{\xi}^{1} \{\mathbb{K}_n(t)\}^2 \, \mathrm{d}t,$$

which is a continuous functional of the process

$$\mathbb{K}_n(t) = \sqrt{n}\,\{K_n(t) - K(\theta_n, t)\}. \tag{2.1}$$

In order to avoid technical difficulties related to censorship and possible unboundedness of the density of $K(\theta, \cdot)$ at the origin, their statistic involves an arbitrary cut-off point $\xi > 0$. Mimicking the approach of Barbe et al. (1996), they were able to identify the limit of $\mathbb{K}_n$, and hence that of $S_{\xi n}$, even under the presence of censoring. However, because of an observed bias in a parametric bootstrap procedure they describe for approximating the variance of

$S_{\xi n}$, Wang & Wells (2000) ended up recommending that the selection of a model from a set of Archimedean copula families be based on a comparison of the raw values of $S_{\xi n}$.

This paper extends the work of Wang & Wells (2000) in a number of ways. Expressed in the simplest of terms, what is proposed here are alternatives to $S_{\xi n}$ given by

$$S_n = \int_0^1 |\mathbb{K}_n(t)|^2 \, k(\theta_n, t) \, \mathrm{d}t \quad \text{and} \quad T_n = \sup_{0 \le t \le 1} |\mathbb{K}_n(t)| \, .$$

It will be seen that the use of these statistics has several advantages. Specifically:

a) simple formulas are available for S_n and T_n in terms of the ranks of the observations, which is not the case for $S_{\xi n}$;

b) the procedures are free of any arbitrary constant ξ, whose selection and influence on the limiting distribution of $S_{\xi n}$ were not addressed by Wang & Wells (2000);

c) the large-sample distribution of S_n and T_n can be found not only for bivariate Archimedean copulas, but in arbitrary dimension $d \geq 2$ and for general copulas satisfying weak regularity conditions;

d) although the limits are not explicit, a parametric bootstrap procedure which is demonstrably valid can be used to approximate P-values associated with any continuous functional of $\mathbb{K}_n$, and in particular with the goodness-of-fit statistics S_n and T_n.

In the course of these developments, an explanation will be given for the bias observed by Wang & Wells (2000) in their own parametric bootstrap, and a correction will be proposed. Furthermore, numerical examples will illustrate how a selection procedure based only on the comparison of raw values of $S_{\xi n}$ may sometimes lead to models that should be rejected on the basis of their P-value. Of course, comparing raw values of S_n and T_n could be just as misleading, whence the importance of the valid parametric bootstrap procedure proposed herein.

Conditions which ensure the convergence of $\mathbb{K}_n$ in general dimension $d \geq 2$ are given in Section 2.2 and verified in Section 2.3 for a number of common copula families, including non-Archimedean models. Bootstrap methods to compute P-values are discussed in Section 2.4, and in Section 2.5, simulations are then used to assess the power of goodness-of-fit tests based on statistics S_n and T_n. Comparisons are also made there with a statistic proposed by Shih (1998) for testing the adequacy of the Clayton family of copulas, also known as the gamma frailty model. Two concrete examples of application of the new procedures are discussed in Section 2.6, and perspectives for future work are highlighted in the final section.

While the work of Wang & Wells (2000) was motivated by biostatistical applications in which data are often censored, the present paper does not address the issue of censorship, as it arose from modelling issues in actuarial science and finance, where this problem is much less frequent. For illustrations of copula modelling in the latter fields, see for instance Frees & Valdez (1998), Klugman & Parsa (1999), Li (2000), Belguise & Lévi (2001–2002), Cherubini et al. (2004), Embrechts et al. (2002), Hennessy & Lapan (2002), Lauprete et al. (2002), Dakhli (2004), van den Goorbergh et al. (2005).

2.2 Distributional results

Let $(X_{11},\ldots,X_{d1}),\ldots,(X_{1n},\ldots,X_{dn})$ be $n \geq 2$ independent copies of a vector $\mathbf{X} = (X_1,\ldots,X_d)$ from some continuous d–variate copula model $\mathcal{C} = (C_\theta)$ with unknown continuous margins $F_1,\ldots,F_d$. In other words, suppose that the cumulative distribution function H of $\mathbf{X}$ is of the form

$$H(x_1,\ldots,x_d) = C\left\{F_1(x_1),\ldots,F_d(x_d)\right\}, \tag{2.2}$$

for some $C = C_\theta \in \mathcal{C}$, whose parameter θ takes its value in an open set $\mathcal{O} \subset \mathbb{R}^m$. It is not assumed that C_θ is Archimedean in the sequel.

Let $K(\theta,t) = P\{H(\mathbf{X}) \leq t\}$, and define its empirical version as

$$K_n(t) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}(V_{jn} \leq t), \quad t \in [0,1] \tag{2.3}$$

where the V_{jn} are pseudo-observations defined by

$$\begin{aligned} V_{jn} &= \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}\left(X_{1k} \leq X_{1j},\ldots,X_{dk} \leq X_{dj}\right) \\ &= \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}\left(R_{1k} \leq R_{1j},\ldots,R_{dk} \leq R_{dj}\right), \end{aligned}$$

and where R_{ij} is the rank of X_{ij} among $X_{i1},\ldots,X_{in}$.

Proposition (2.1) below identifies the limiting distribution $\mathbb{K}$ of the process $\mathbb{K}_n$ defined in (2.1). The conditions under which this convergence occurs, stated as hypotheses I–IV, are similar to those of Theorem 1 of Barbe et al. (1996), who establish the asymptotic behavior of Kendall's process, defined as

$$\mathbb{K}_{n,\theta}(t) = \sqrt{n}\left\{K_n(t) - K(\theta, t)\right\}. \tag{2.4}$$

Under these hypotheses, the statistics S_n and T_n are continuous functionals of $\mathbb{K}_n$. Their limits are thus given respectively by

$$S = \int_0^1 |\mathbb{K}(t)|^2 \, k(\theta, t)\, \mathrm{d}t \quad \text{and} \quad T = \sup_{t \in [0,1]} |\mathbb{K}(t)|\,. \tag{2.5}$$

Hypothesis I

For all $\theta \in \mathcal{O}$, the distribution function $K(\theta, t)$ of $H(\mathbf{X})$ admits a density $k(\theta, t)$ which is continuous on $\mathcal{O} \times (0, 1]$ and such that

$$k(\theta, t) = o\left\{t^{-1/2} \log^{-1/2-\epsilon}(1/t)\right\}$$

for some $\epsilon > 0$ as $t \to 0$.

Hypothesis II

For all $\theta \in \mathcal{O}$, there exists a version of the conditional distribution of the vector $\mathbf{X} = (X_1, \ldots, X_d)$ given $H(\mathbf{X}) = t$ such that, for any continuous real-valued function f on $[0, 1]^d$, the mapping

$$t \mapsto \mu(t, f) = k(\theta, t)\ \mathrm{E}\left\{f\left(X_1, \ldots, X_d\right) \mid H(\mathbf{X}) = t\right\}$$

is continuous on $(0, 1]$ with $\mu(1, f) = k(\theta, 1)\ f(1, \ldots, 1)$.

Let $\Theta_n = \sqrt{n}\,(\theta_n - \theta)$. The next hypothesis merely requires θ_n to be a “good” estimator of the parameter θ.

Hypothesis III

The sequence of vectors $(\mathbb{K}_{n,\theta}, \Theta_n)$ *converges in law to a centered Gaussian vector* $(\mathbb{K}_\theta, \Theta)$ *with limiting covariance*

$$\gamma(t) = (\gamma_1(t), \ldots, \gamma_m(t))^\top = (\mathrm{cov}\left\{\mathbb{K}_\theta(t), \Theta_1\right\}, \ldots, \mathrm{cov}\left\{\mathbb{K}_\theta(t), \Theta_m\right\})^\top,$$

where $t \in [0,1]$ *and* $\mathrm{var}(\Theta) = \Sigma$.

Before stating hypothesis IV, define

$$\dot{K}(\theta, t) = \nabla_\theta K(\theta, t) = \left(\frac{\partial}{\partial\theta_1} K(\theta, t), \ldots, \frac{\partial}{\partial\theta_m} K(\theta, t)\right)^\top$$

to be the gradient of $K(\theta, t)$ with respect to θ.

Hypothesis IV

For every given $\theta \in \mathcal{O}$, $\dot{K}(\theta, t)$ *exists and is continuous for all* $t \in [0,1]$. *Moreover,*

$$\sup_{\|\theta^\star - \theta\| < \varepsilon} \sup_{t \in [0,1]} \left| \dot{K}(\theta^\star, t) - \dot{K}(\theta, t) \right| \longrightarrow 0 \quad as \quad \varepsilon \to 0. \tag{2.6}$$

Remark 2.1. *If* $\ddot{K}(\theta, t) = \nabla_\theta \dot{K}(\theta, t)$, *then condition (2.6) is verified whenever*

$$\sup_{\|\theta^\star - \theta\| < \varepsilon} \sup_{t \in [0,1]} \left\| \ddot{K}(\theta^\star, t) \right\| < \infty, \quad as\ \varepsilon > 0.$$

As noted by Ghoudi & Rémillard (1998, 2004), who studied the asymptotic behavior of empirical processes constructed from the more general concept of pseudo-observation, hypotheses I and II imply the weak convergence of the empirical process (2.4) to a Gaussian limit $\mathbb{K}_\theta$ having zero mean. From the work of Barbe et al. (1996), the covariance function of $\mathbb{K}_\theta$ is of the form

$$\begin{aligned}\Gamma_\theta(s,t) &= K(\theta, s\wedge t) - K(\theta,s)K(\theta,t)\\ &\quad + k(\theta,s)k(\theta,t)R_\theta(s,t) - k(\theta,t)Q_\theta(s,t) - k(\theta,s)Q_\theta(t,s),\end{aligned}$$

where $s\wedge t = \min(s,t)$ and for all $s,t\in[0,1]$,

$$R_\theta(s,t) = \mathrm{P}\{\mathbf{X}_1 \le \mathbf{X}_2 \wedge \mathbf{X}_3 \mid H(\mathbf{X}_2)=s, H(\mathbf{X}_3)=t\} - st$$

and

$$Q_\theta(s,t) = \mathrm{P}\{H(\mathbf{X}_1)\le s, \mathbf{X}_1 \le \mathbf{X}_2 \mid H(\mathbf{X}_2)=t\} - tK(\theta,s)$$

are defined in terms of mutually independent copies $\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{X}_3$ of $\mathbf{X}$.

As stated next, hypotheses III and IV allow one to obtain a similar result for the empirical process $\mathbb{K}_n$.

Proposition 2.1. *Under hypotheses I–IV, the empirical process* $\mathbb{K}_n(t) = \sqrt{n}\,\{K_n(t) - K(\theta_n,t)\}$ *converges in* $\mathcal{D}[0,1]$ *to a continuous, centered Gaussian process having representation*

$$\mathbb{K}(t) = \mathbb{K}_\theta(t) - \dot{K}(\theta,t)^\top\Theta, \quad t\in[0,1]$$

and covariance function

$$\Gamma(s,t) = \Gamma_\theta(s,t) + \dot{K}(\theta,s)^\top\Sigma\ \dot{K}(\theta,t) - \dot{K}(\theta,s)^\top\gamma(t) - \dot{K}(\theta,t)^\top\gamma(s)$$

for all $s,t\in[0,1]$.

Proof. Write $\mathbb{K}_n(t) = \mathbb{K}_{n,\theta}(t) - B_n(t)$ with

$$B_n(t) = \sqrt{n}\,\{K(\theta_n, t) - K(\theta, t)\}$$

for all $t \in [0,1]$. In view of hypotheses I and II, $\mathbb{K}_{n,\theta}$ converges in $\mathcal{D}[0,1]$ to a continuous, centered Gaussian process $\mathbb{K}_\theta$ with covariance function Γ_θ, as defined above. Furthermore, it is shown in Section 1 of Appendix A that

$$\sup_{t\in[0,1]} \left| B_n(t) - \dot{K}(\theta, t)^\top \Theta_n \right| \xrightarrow{P} 0$$

under hypotheses III and IV. Finally, making use of hypothesis III, one has

$$\begin{aligned}
\Gamma(s,t) &= \operatorname{cov}\{\mathbb{K}(s), \mathbb{K}(t)\} \\
&= \operatorname{cov}\{\mathbb{K}_\theta(s), \mathbb{K}_\theta(t)\} + \operatorname{cov}\left\{\dot{K}(\theta, s)^\top\Theta,\ \dot{K}(\theta, t)^\top\Theta\right\} \\
&\quad - \operatorname{cov}\left\{\dot{K}(\theta, s)^\top\Theta,\ \mathbb{K}_\theta(t)\right\} - \operatorname{cov}\left\{\dot{K}(\theta, t)^\top\Theta,\ \mathbb{K}_\theta(s)\right\} \\
&= \Gamma_\theta(s,t) + \dot{K}(\theta, s)^\top\Sigma\ \dot{K}(\theta, t) - \dot{K}(\theta, s)^\top\gamma(t) - \dot{K}(\theta, t)^\top\gamma(s).
\end{aligned}$$

Thus the proof is complete. ◊

A potential candidate for θ_n is the omnibus rank-based estimator of Genest et al. (1995) or Shih & Louis (1996a), obtained by maximizing the pseudo-likelihood

$$\sum_{j=1}^{n} \log c_\theta\left(\frac{R_{1j}}{n+1}, \ldots, \frac{R_{dj}}{n+1}\right),$$

where c_θ is the density associated with C_θ. It follows from example 3.2.1 of Ghoudi & Rémillard (2004) that hypothesis III is automatically verified when hypotheses I–II hold, so that proposition (2.1) holds under hypotheses I, II, and IV only.

In the special case where θ is real, another procedure considered by Wang & Wells (2000), among others, consists of estimating θ by $\theta_n = \tau^{-1}(\tau_n)$, where $\tau(\theta)$ is the multivariate extension of Kendall's tau defined by

$$\tau = \left(\frac{2^d - 1}{2^{d-1} - 1}\right) - \left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 K(\theta, t)\, \mathrm{d}t, \tag{2.7}$$

as in Barbe et al. (1996) or Jouini & Clemen (1996). Assuming that this mapping has a continuous non-vanishing derivative

$$\dot{\tau}(\theta) = -\left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 \dot{K}(\theta, t)\, \mathrm{d}t$$

in $\mathcal{O}$. Letting

$$\tau_n = \left(\frac{2^d - 1}{2^{d-1} - 1}\right) - \left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 K_n(t)\, \mathrm{d}t, \tag{2.8}$$

one can see that $\sqrt{n}\,(\tau_n - \tau)$ is related to Kendall's process $\mathbb{K}_{n,\theta}$ through the linear functional

$$\sqrt{n}\,(\tau_n - \tau) = -\left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 \mathbb{K}_{n,\theta}(t)\, \mathrm{d}t.$$

An application of Slutsky's theorem then implies that, under hypotheses I–II,

$$\Theta_n = \sqrt{n}\left\{\tau^{-1}(\tau_n) - \theta\right\} = \frac{-1}{\dot{\tau}(\theta)} \left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 \mathbb{K}_{n,\theta}(t)\, \mathrm{d}t + o_P(1)$$

converges in law to

$$\Theta = \frac{-1}{\dot{\tau}(\theta)} \left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 \mathbb{K}_\theta(t)\, \mathrm{d}t = \kappa_\theta \int_0^1 \mathbb{K}_\theta(t)\, \mathrm{d}t,$$

where

$$1/\kappa_\theta = \int_0^1 \dot{K}(\theta, t)\, \mathrm{d}t$$

is assumed to be nonzero. The joint convergence of $(\mathbb{K}_{n,\theta}, \Theta_n)$ to $(\mathbb{K}_\theta, \Theta)$ required in hypothesis III is thus immediate, and

$$\gamma(t) \quad = \quad \operatorname{cov}\left\{\mathbb{K}_\theta(t), \Theta\right\} = \kappa_\theta \int_0^1 \Gamma_\theta(u,t)\,\mathrm{d}u$$

while

$$\operatorname{var}(\Theta) = \kappa_\theta^2 \int_0^1 \int_0^1 \Gamma_\theta(u,v)\,\mathrm{d}u\,\mathrm{d}v.$$

This suggests the following consequence of the main result.

Proposition 2.2. *If $\theta \in \mathcal{O} \subseteq \mathbb{R}$ is estimated by $\theta_n = \tau^{-1}(\tau_n)$ and κ_θ is finite, then under hypotheses I, II and IV, the empirical process $\mathbb{K}_n$ converges weakly to a centered Gaussian process having representation*

$$\mathbb{K}(t) = \mathbb{K}_\theta(t) - \kappa_\theta \dot{K}(\theta,t) \int_0^1 \mathbb{K}_\theta(v)\,\mathrm{d}v, \quad t \in [0,1]$$

and limiting covariance function

$$\begin{aligned}\Gamma(s,t) \quad = \quad & \Gamma_\theta(s,t) + \kappa_\theta^2 \dot{K}(\theta,s)\dot{K}(\theta,t) \int_0^1 \int_0^1 \Gamma_\theta(u,v)\,\mathrm{d}u\,\mathrm{d}v \\ & - \kappa_\theta \dot{K}(\theta,s) \int_0^1 \Gamma_\theta(u,t)\,\mathrm{d}u - \kappa_\theta \dot{K}(\theta,t) \int_0^1 \Gamma_\theta(s,v)\,\mathrm{d}v\end{aligned}$$

for all $s,t \in [0,1]$.

2.3 Examples

This section presents a few popular classes of multivariate copulas that satisfy hypotheses I–IV stated above. The list is by no means exhaustive, of course.

2.3.1 Archimedean copulas

Copulas are called Archimedean when they may be expressed in the form

$$C(u_1, \ldots, u_d) = \phi^{-1}\left\{\phi(u_1) + \cdots + \phi(u_d)\right\} \tag{2.9}$$

in terms of a bijection $\phi : (0, 1] \to [0, \infty)$ such that $\phi(1) = 0$ and

$$(-1)^i \frac{\mathrm{d}^i}{\mathrm{d}x^i}\,\phi^{-1}(x) > 0, \quad i \in \{1, \ldots, d\}. \tag{2.10}$$

As shown by Genest & Rivest (1993) in the case $d = 2$, the generator ϕ can be recovered from K, since

$$K(t) = t - \frac{\phi(t)}{\phi'(t)}, \quad t \in (0, 1].$$

Among the multivariate copula models that fall into this category (see Nelsen, 1999, chapter 4), Table 2.1 presents summary information for those of Ali et al. (1978), Clayton (1978), Gumbel (1960) and Frank (1979). Note that in this table, the parameter space $\mathcal{O}$ is limited to positive degrees of association, as those are the only values that can be achieved in all dimensions for Archimedean copulas in general (Marshall & Olkin, 1988), and for these four models in particular.

One key characteristic of Archimedean copulas is the fact that all the information about the $d-$dimensional dependence structure is contained in a univariate generator, ϕ_θ. From Barbe et al. (1996),

$$K(\theta, t) = t + \sum_{i=1}^{d-1} \frac{(-1)^i}{i!} \{\phi_\theta(t)\}^i f_i(\theta, t), \tag{2.11}$$

where

$$f_i(\theta, t) = \left. \frac{\mathrm{d}^i}{\mathrm{d}x^i} \phi_\theta^{-1}(x) \right|_{x=\phi_\theta(t)},$$

provided that $\{\phi_\theta(t)\}^i f_i(\theta, t) \to 0$ as $t \to 0$ for all $i \in \{1, \ldots, d-1\}$. Note in passing that

$$f_{i+1}(\theta, t) = f_1(\theta, t)\, \partial f_i(\theta, t)/\partial t, \quad i \in \{1, \ldots, d-1\}. \tag{2.12}$$

As a consequence, the moments of C_θ are also functions of ϕ_θ only. In addition, the multivariate version of Kendall's measure of association may be computed, in view of (2.7), through the formula

$$\tau = 1 - \left(\frac{2^d}{2^{d-1} - 1} \right) \sum_{i=1}^{d-1} \frac{(-1)^i}{i!} \int_0^1 \{\phi_\theta(t)\}^i f_i(\theta, t)\, \mathrm{d}t,$$

which for $d = 2$ reduces to the well-known expression (Genest & MacKay, 1986; Nelsen, 1999, section 5.1)

$$\tau = 1 + 4 \int_0^1 \frac{\phi_\theta(t)}{\phi_\theta'(t)}\, \mathrm{d}t.$$

Barbe et al. (1996) prove that the four families of copulas listed in Table 2.1 meet hypotheses I and II. It is shown in Sections 2.1–2.4 of Appendix A that they also satisfy hypothesis IV for all values of $\theta \in \mathcal{O}$. Explicit expressions for K may also be found there.

Table 2.1: Families of multivariate Archimedean copulas

Model	$\phi_\theta(t)$	$K(\theta,t)$ for $d=2$	$\tau = g(\theta)$	$\mathcal{O}$
Ali–Mikhail–Haq	$\frac{\log\left(\frac{1-\theta}{t}+\theta\right)}{1-\theta}$	$t+\frac{t^2}{1-\theta}\left(\frac{1-\theta}{t}+\theta\right)\times \log\left(\frac{1-\theta}{t}+\theta\right)$	$\frac{3\theta-2}{3\theta} - \frac{2(1-\theta)^2\log(1-\theta)}{3\theta^2}$	$(0,1)$
Clayton	$(t^{-\theta}-1)/\theta$	$t+t(1-t^{\theta})/\theta$	$\theta/(\theta+2)$	$(0,\infty)$
Frank	$\log\left(\frac{1-e^{-\theta}}{1-e^{-\theta t}}\right)$	$t-\frac{(1-e^{\theta t})}{\theta}\log\left(\frac{1-e^{-\theta}}{1-e^{-\theta t}}\right)$	$1-\frac{4}{\theta}+\frac{4D_1(\theta)}{\theta}$	$(0,\infty)$
Gumbel–Hougaard	$\lvert\log t\rvert^{1/(1-\theta)}$	$t-(1-\theta)t\log t$	θ	$(0,1)$

Here, $D_1(\theta) = \theta^{-1}\int_0^\theta \frac{x}{e^x-1}\,\mathrm{d}x$ stands for the first Debye function

2.3.2 Bivariate extreme-value copulas

It has been known since the work of Pickands (1981) that bivariate extreme-value distributions have underlying copulas of the form

$$C_A(u_1, u_2) = \exp\left[\log(u_1 u_2) A\left\{\frac{\log(u_1)}{\log(u_1 u_2)}\right\}\right],$$

where the dependence function A, defined on $[0, 1]$, is convex and such that

$$\max(t, 1-t) \leq A(t) \leq 1$$

for all $t \in [0, 1]$. The most common parametric models of bivariate extreme-value copulas are presented in Table 2.2. For additional details see, for instance, Tawn (1988), Capéraà et al. (1997); Capéraà & Fougères (2001) or Capéraà & Fougères (2001).

Ghoudi et al. (1998) note that for a random vector (U_1, U_2) distributed as C_A, then

$$K_A(t) = P\left\{C_A(U_1, U_2) \leq t\right\} = t - (1-\tau)t\log t, \quad t \in (0, 1]$$

depends only on the population version of Kendall's measure of association computed as a function of A through the identity

$$\tau = \tau(A) = \int_0^1 \frac{t(1-t)}{A(t)}\, \mathrm{d}A'(t).$$

Thus if two bivariate extreme-value copulas with generators A and A^* are such that $\tau(A) = \tau(A^*)$, then $K_A = K_{A^*}$ and could not possibly be distinguished by goodness-of-fit procedures based on the process $\mathbb{K}_n$.

Table 2.2: Families of bivariate extreme-value copulas

Model	$A_\theta(t)$	$C_{A_\theta}(u_1, u_2)$	$\mathcal{O}$
Gumbel	$\theta t^2 - \theta t + 1$	$u_1 u_2 \exp\left(-\theta \dfrac{\log u_1 \log u_2}{\log u_1 u_2}\right)$	$(0, 1)$
Gumbel–Hougaard	$\left\{t^{\frac{1}{1-\theta}} + (1-t)^{\frac{1}{1-\theta}}\right\}^{1-\theta}$	$\exp\left[-\left\{\lvert\log u_1\rvert^{\frac{1}{1-\theta}} + \lvert\log u_2\rvert^{\frac{1}{1-\theta}}\right\}^{1-\theta}\right]$	$(0, 1)$
Galambos	$1 - \left\{t^{-\theta} + (1-t)^{-\theta}\right\}^{-1/\theta}$	$u_1 u_2 \exp\left[\left(\lvert\log u_1\rvert^{-\theta} + \lvert\log u_2\rvert^{-\theta}\right)^{-1/\theta}\right]$	$(0, \infty)$
MO	$\max\{1 - \theta_1 t, 1 - \theta_2(1-t)\}$	$u_1^{1-\theta_1} u_2^{1-\theta_2} \min\left(u_1^{\theta_1}, u_2^{\theta_2}\right)$	$(0, 1)^2$

Whatever the case may be, it can be checked easily that the conditions of the main result are satisfied for this large class of models. To this end, first note that $\dot{K}(\tau, t) = t \log t$, so that hypothesis IV is satisfied. As hypothesis I is also trivially verified, an application of proposition 2.2 implies that

$$\mathbb{K}_n(t) = \sqrt{n}\,\{K_n(t) - K(\tau_n, t)\}$$

converges to

$$\mathbb{K}(t) = \mathbb{K}_\tau(t) + 4\, t \log t \int_0^1 \mathbb{K}_\tau(v)\, \mathrm{d}v.$$

Furthermore, the limiting covariance function, for which no explicit representation seems possible, is given by

$$\begin{aligned}\Gamma(s,t) &= \Gamma_\tau(s,t) + 16st\log s\log t\int_0^1\int_0^1 \Gamma_\tau(u,v)\,\mathrm{d}u\,\mathrm{d}v \\ &\quad + 4s\log s\int_0^1 \Gamma_\tau(u,t)\,\mathrm{d}u + 4t\log t\int_0^1 \Gamma_\tau(s,v)\,\mathrm{d}v.\end{aligned}$$

2.3.3 Fréchet copulas

These bivariate copulas are mixtures of the independence copula $C_\Pi(u,v) = uv$ and of the upper Fréchet–Hoeffding bound $C_M(u,v) = \min(u,v)$, that is,

$$C_\theta(u,v) = (1-\theta)uv + \theta\min(u,v), \quad \theta \in [0,1].$$

Letting $\zeta(\theta,t) = 4t/\{I(\theta,t)+\theta\}^2$ and $I(\theta,t) = \{\theta^2 + 4t(1-\theta)\}^{1/2}$, one can show (see Genest & Rivest 2001 for a proof) that

$$K(\theta,t) = t - t\log t + t\log\{\zeta(\theta,t)\}, \quad t \in [0,1].$$

Note that ζ is continuous on $[0,1]^2$ and bounded above by 1.

For this model, it is known (see, for example, Nelsen 1999, p. 130) that $\tau = \theta(\theta+2)/3$, and hence $\theta = \sqrt{3\tau+1} - 1$. Furthermore, a simple calculation shows that $\zeta'(\theta,t)/\zeta(\theta,t) = \theta/\{tI(\theta,t)\}$, whence the density associated with $K(\theta,t)$ is given by $k(\theta,t) = -\log t + \log\{\zeta(\theta,t)\} + \theta/I(\theta,t)$. The latter function is continuous and since $\theta \le I(\theta,t)$, one can see at once that it has the appropriate behavior as $t \to 0$ for hypothesis I to hold. The verification of hypothesis IV is deferred to Section 2.5 of Appendix A.

2.3.4 Bivariate Farlie–Gumbel–Morgenstern copulas

It is not always necessary to be able to compute $K(\theta, t)$ explicitly to obtain the weak convergence of $\mathbb{K}_{n,\theta}$. Such is the case for the Farlie–Gumbel–Morgenstern system of distributions, whose members have the form

$$C_\theta(u, v) = uv + \theta uv(1-u)(1-v), \quad \theta \in [-1, 1].$$

Barbe et al. (1996) show that

$$k(\theta, t) = \int_t^1 h(\theta, x, t)\,\mathrm{d}x,$$

where

$$h(\theta, x, t) = \left\{ \frac{1}{(1-x)r(\theta, x, t)} + \frac{1}{x} - \frac{1}{1-x} \right\} \mathbf{1}\,(t < x \leq 1)$$

and

$$r(\theta, x, t) = \left[\{1 - \theta(1-x)\}^2 + 4\theta(1-x)(1-t/x) \right]^{1/2}.$$

Although in this model $\tau = 2\theta/9$, there is no explicit expression for

$$K(\theta, t) = \int_0^t k(\theta, s)\,\mathrm{d}s = \int_0^t \int_s^1 h(\theta, x, s)\,\mathrm{d}x\,\mathrm{d}s.$$

Nevertheless, Barbe et al. (1996) prove that $k(\theta, t)$ satisfies hypothesis I. They further mention that hypothesis II is verified as well. The proof that hypothesis IV also holds may be found in Appendix A, Section 2.6.

2.4 Implementation of the goodness-of-fit tests

Straightforward calculations show that

$$\begin{aligned} S_n &= \frac{n}{3} + n \sum_{j=1}^{n-1} K_n^2\left(\frac{j}{n}\right)\left\{K\left(\theta_n, \frac{j+1}{n}\right) - K\left(\theta_n, \frac{j}{n}\right)\right\} \\ &\quad - n \sum_{j=1}^{n-1} K_n\left(\frac{j}{n}\right)\left\{K^2\left(\theta_n, \frac{j+1}{n}\right) - K^2\left(\theta_n, \frac{j}{n}\right)\right\} \end{aligned}$$

and

$$T_n = \sqrt{n} \max_{i=0,1;\, 0 \le j \le n-1} \left\{\left|K_n\left(\frac{j}{n}\right) - K\left(\theta_n, \frac{j+i}{n}\right)\right|\right\}.$$

Formal testing procedures based on these statistics would consist of rejecting $H_0 : C \in \mathcal{C}$ when the observed value of S_n or T_n is greater than the $100(1-\alpha)\%$ percentile of its distribution under the null hypothesis. As implied by formula (2.5), however, this distribution depends on the unknown association parameter θ, even in the limit.

This fact is illustrated in Table 2.3, where the 95th percentiles of the distributions of S_n and T_n are evaluated for Archimedean copulas of Table 1 for some values of τ. Of course, the $S_{\xi n}(\theta_n)$ statistic of Wang & Wells (2000) suffers from the same limitation, let alone its dependence on the arbitrary cut-off point ξ.

To circumvent these methodological issues and obtain an approximate P-value for either S_n or T_n, one may call on a parametric bootstrap or Monte Carlo testing approach based on C_{θ_n}. This is described is Section 2.4.2. Since the tests and their distributions only involve C_{θ_n} through $K(\theta_n, \cdot)$, it may be tempting to base the bootstrap on the latter only, as done by Wang & Wells (2000). Section 2.4.1 explains why this shortcut is inappropriate.

Table 2.3: Estimation based on 1000 replicates of the 95th percentile of the distribution of the Cramér–von Mises statistic S_n and the Kolmogorov–Smirnov statistic T_n

		S_n			T_n		
Model	τ	$n = 100$	$n = 250$	$n = 1000$	$n = 100$	$n = 250$	$n = 1000$
Clayton	0.20	0.1872	0.1589	0.1567	1.0402	0.9725	0.9824
	0.40	0.1410	0.1336	0.1278	0.9244	0.9015	0.8805
	0.60	0.0992	0.0902	0.0977	0.8563	0.8071	0.7863
	0.80	0.0573	0.0518	0.0511	0.6775	0.6578	0.6650
Frank	0.20	0.1515	0.1328	0.1216	0.9743	0.9294	0.9210
	0.40	0.1254	0.1186	0.1150	0.8883	0.8614	0.8439
	0.60	0.1017	0.0979	0.0975	0.7945	0.7982	0.7875
	0.80	0.0591	0.0536	0.0492	0.6410	0.6472	0.6496
Gumbel–Hougaard	0.20	0.1516	0.1266	0.1265	0.9740	0.9431	0.9325
	0.40	0.1255	0.1155	0.1117	0.9284	0.8812	0.8737
	0.60	0.0946	0.0961	0.0833	0.8033	0.7839	0.8000
	0.80	0.0567	0.0550	0.0508	0.6381	0.6573	0.6469

2.4.1 Parametric bootstrap method suggested by Wang & Wells (2000)

To find an estimate of the variance of their statistic $S_{\xi n}$, Wang & Wells (2000) propose to generate a sample $V_{1,n}^*, \dots, V_{n,n}^*$, where $V_{j,n}^* \sim K(\theta_n, \cdot)$ for $j \in \{1, \dots, n\}$, and then to calculate

$$K_n^*(v) = \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}\left(V_{j,n}^* \leq v\right),$$

$$\tau_n^* = -1 + \frac{4}{n} \sum_{j=1}^{n} V_{j,n}^*, \quad \theta_n^* = \tau^{-1}\left(\tau_n^*\right)$$

and

$$S_{\xi n}^* = n \int_{\xi}^{1} \left\{K_n^*(v) - K\left(\theta_n^*, v\right)\right\}^2 \, \mathrm{d}v.$$

By repeating the procedure N times, one ends up with values $S_{\xi n,1}^*, \dots, S_{\xi n,N}^*$. Wang & Wells (2000) thus propose to estimate the variance of $S_{\xi n}$ by the sample variance of $S_{\xi n,1}^*, \dots, S_{\xi n,N}^*$.

Unfortunately, this procedure is invalid. As shown in Section 3 of Appendix A, the empirical bootstrap process $\sqrt{n}\{\mathbb{K}_n^* - K(\theta_n^*, \cdot)\}$ converges in $\mathcal{D}[0,1]$ to a process $\mathbb{K}^*$ which is independent of $\mathbb{K}$ but generally different in law. Consequently, the proposed bootstrap procedure does not yield a valid estimate of the variance of $S_{\xi n}$, nor reliable P-values of any goodness-of-fit procedure based thereon.

2.4.2 Parametric bootstrap method based on C_{θ_n}

In order to compute P-values for any test statistic based on the empirical process $\mathbb{K}_n$, one requires generating a large number, N, of independent samples of size n from C_{θ_n} and computing the corresponding values of the selected statistic, such as S_n or T_n. In the former case, for example, the procedure would work as follows:

Step 1: Estimate θ by a consistent estimator θ_n.

Step 2: Generate N random samples of size n from C_{θ_n} and, for each of these samples, estimate θ by the same method as before and determine the value of the test statistic.

Step 3: If $S^*_{1:N} \leq \cdots \leq S^*_{N:N}$ denote the ordered values of the test statistics calculated in Step 2, an estimate of the critical value of the test at level α based on S_n is given by $S^*_{\lfloor(1-\alpha)N\rfloor:N}$, and

$$\frac{1}{N} \# \left\{ j : S^*_j \geq S_n \right\}$$

yields an estimate of the P-value associated with the observed value S_n of the statistic.

The validity of this approach is established in a companion paper by Genest & Rémillard (2005). The assumptions needed for the method to work are stated in Section 4 of Appendix A.

2.5 Numerical studies

Simulation studies were conducted to assess the finite-sample properties of the proposed goodness-of-fit tests for various classes of copula models under the null hypothesis and under the alternative. Three copula families were used under H_0, namely those of Clayton, Frank, and Gumbel–Hougaard. All of them are Archimedean and complete, in the sense that they cover all possible degrees of positive dependence, as measured by Kendall's tau. Three complete systems of non-Archimedean copulas were also used as alternatives, namely the Fréchet, Gaussian, and Plackett families. In all models considered, the validity conditions for the parametric bootstrap are verified.

In each case, 10,000 pseudo-random samples of size $n = 250$ were generated from the selected model with a specified value of Kendall's tau, chosen in the set $\{0.2, 0.4, 0.6, 0.8\}$. For each sample, the dependence parameter of the copula model under the null hypothesis was estimated by inversion of Kendall's tau, that is, by setting $\theta_n = g^{-1}(\tau_n)$ for the appropriate function g. For the Clayton and Gumbel–Hougaard copulas, the estimators were thus simply

$$\theta_n = \frac{2\tau_n}{1-\tau_n} \quad \text{and} \quad \theta_n = \tau_n,$$

respectively, but the inversion needed to be carried out numerically for Frank's model. Table 2.4 shows that the resulting rank-based estimators have reasonably small percent relative bias for samples of size $n = 100$ and 250.

Table 2.4: Percent relative bias of the estimator $\theta_n = g^{-1}(\tau_n)$ based on 10,000 samples of size $n = 100$ and $n = 250$ from the Clayton, Frank and Gumbel–Hougaard models with various degrees of dependence

Model	τ	$n = 100$	$n = 250$
Clayton	0.20	3.22	1.98
	0.40	3.25	0.95
	0.60	2.39	0.93
	0.80	2.29	0.98
Frank	0.20	1.57	0.56
	0.40	1.26	0.23
	0.60	1.10	0.44
	0.80	1.28	0.28
Gumbel–	0.20	0.12	0.18
Hougaard	0.40	0.07	0.01
	0.60	−0.05	0.05
	0.80	0.01	0.01

2.5.1 Comparison between S_n, T_n and $S_{\xi n}$

Table 2.5 shows the power and size of S_n and T_n as goodness-of-fit test statistics of the Clayton, Frank, and Gumbel–Hougaard null hypothesis under 24 choices of copula for the true distribution. Results for the $S_{\xi n}$ statistic of Wang & Wells (2000) with $\xi = 0$ were also added for comparison purposes. In interpreting these results, it must be kept in mind that the error associated with the power estimates is larger than might usually be expected under 10,000 replications. This is because the distribution of the observations under the null hypothesis involves a parameter that must be estimated. Additional variation thus arises from the use of C_{θ_n} rather than C_θ in the calculation of the P-values associated with the tests.

Table 2.5 shows that when the null hypothesis holds true, all three statistics are at the right level, up to sampling error. Detailed inspection of the results leads to the following additional insights:

a) As a general rule, the tests S_n and S_{0n} based on Cramér–von Mises functionals outperform that which is founded on the Kolmogorov–Smirnov statistic T_n.

b) All tests appear to distinguish rather easily between the Clayton model and the alternatives considered; the best performance is achieved by S_{0n}.

c) Testing for the Frank or the Gumbel–Hougaard families seems to be more difficult, at least given the alternatives considered; each of S_n and S_{0n} delivers the best power in roughly half the cases.

There is also a hint in Table 2.5 that as the value of τ goes from 0 to 1, the power of the three tests increases, levels off, and ultimately starts decreasing again. This was only to be expected, as all families considered have independence and the Fréchet–Hoeffding upper bound $M(u, v) = u \wedge v$ as their limiting copulas at $\tau = 0$ and $\tau = 1$, respectively.

Table 2.5: Percentage of rejection of three different null hypotheses using S_n, T_n or S_{0n} at the 5% level when $n = 250$, based on 10,000 replicates

Alternative		Model under the null hypothesis								
Family	τ	Clayton			Frank			Gumbel–Hougaard		
		S_n	T_n	S_{0n}	S_n	T_n	S_{0n}	S_n	T_n	S_{0n}
Clayton	.2	4.6	5.5	4.9	82.2	78.9	73.6	95.3	88.4	94.4
	.4	4.2	4.7	3.6	99.8	99.3	99.6	100.0	99.9	100.0
	.6	4.6	4.2	5.0	100.0	99.9	99.9	100.0	100.0	100.0
	.8	4.7	4.4	5.6	100.0	99.8	100.0	100.0	100.0	100.0
Frank	.2	47.1	38.9	51.1	4.7	4.3	5.7	24.5	14.7	30.9
	.4	97.1	88.4	97.4	3.7	5.0	5.9	55.8	37.4	65.6
	.6	99.9	98.2	100.0	5.3	4.9	5.7	75.2	54.2	85.7
	.8	100.0	96.2	100.0	4.4	5.4	3.8	89.1	62.3	88.7
Gumbel–Hougaard	.2	68.4	57.6	80.6	10.6	9.3	25.7	5.1	4.3	5.2
	.4	99.8	97.9	100.0	36.8	24.3	58.4	6.7	5.8	4.3
	.6	100.0	100.0	100.0	69.5	49.5	82.9	4.3	5.4	4.5
	.8	100.0	100.0	100.0	91.2	59.6	89.9	4.4	5.1	4.0
Fréchet	.2	35.0	21.9	41.8	25.3	22.5	20.1	33.1	26.0	22.3
	.4	81.7	60.2	89.2	52.1	38.4	52.2	54.6	40.9	38.1
	.6	98.0	83.1	98.9	77.2	45.5	76.9	51.7	31.2	45.6
	.8	98.3	86.3	99.4	80.1	30.2	78.5	33.9	8.5	28.6
Gaussian	.2	33.3	23.7	38.7	9.9	8.9	9.1	29.2	17.9	24.2
	.4	86.9	71.6	90.7	23.6	17.4	23.5	52.4	34.5	47.5
	.6	99.1	92.1	99.8	50.6	37.7	50.4	53.1	36.5	54.6
	.8	100.0	98.0	100.0	76.8	42.6	75.0	41.3	23.1	40.6
Plackett	.2	49.5	35.1	49.7	5.2	6.4	6.2	20.3	15.7	28.5
	.4	94.0	83.0	95.3	6.3	5.4	7.6	46.5	28.1	48.9
	.6	99.5	94.8	99.8	15.5	10.2	15.7	48.5	31.6	57.5
	.8	99.9	93.5	99.7	36.1	14.6	27.9	40.2	22.9	42.9

2.5.2 Comparison with the test of Shih for the Clayton family

Clayton's copula is often referred to as the gamma frailty model in the survival analysis literature. For this specific choice of null hypothesis, two goodness-of-fit tests are already available, which were developed by Shih (1998) in the bivariate case and by Glidden (1999) for arbitrary dimension $d \geq 2$. It may thus be of interest to compare the power of these specific test statistics to those of the omnibus procedures based on S_n and T_n.

In attempting to make such comparisons, difficulties were encountered with the implementation of both Shih's and Glidden's procedures. Specifically:

a) The limiting variance of the test statistic given on p. 198 of Shih (1998) is erroneous; in fact, her expression tends to $-\infty$ as $\theta = 1/\eta \to 0$, while the correct result is 7/9 in the limiting case of independence. As shown by Genest et al. (2005b), the correct expression for the asymptotic variance should be

$$\frac{18\eta^7 + 240\eta^6 + 3001\eta^5 + 8281\eta^4 + 9449\eta^3 + 5171\eta^2 + 1352\eta + 136}{3\eta^2(\eta+1)^2(3\eta+1)} + \frac{8(2\eta+1)^4}{(\eta+1)^2} L(\eta) - (\eta+1)^4 \left\{\Psi'(1+\eta/2) - \Psi'(1/2+\eta/2)\right\} - 8(\eta+1)(2\eta+1)^2 J(\eta),$$

where

$$
\begin{aligned}
L(\eta) &= \frac{1}{4\eta^2} \text{hypergeom}([1,1,\eta],[2\eta+1,2\eta+1],1), \\
J(\eta) &= \frac{1}{2\eta^2} \text{hypergeom}([1,1,\eta],[\eta+1,2\eta+1],1),
\end{aligned}
$$

and Ψ' denotes the trigamma function.

b) Glidden's test is overly conservative, especially in cases of weak dependence. This phenomenon, documented by Glidden (1999) himself, hints to the fact that the asymptotic distribution may be incorrectly approximated by his proposed computational procedure. In an attempt to reproduce his calculations, it was further discovered that in Glidden's paper, many expressions leading to the identification of the limit were themselves incorrect. In the formulas for ϵ_i and π_k on pp. 385–386, for example, one should replace every instance of $X_{ik\ell}$ by $\tau \wedge X_{ik\ell}$. Furthermore, the expression given for $V(\theta)$ on top of p. 386 should be the limit of $-(1/n)\partial^2 \hat{\ell}_n(\theta)/\partial\theta^2$. Also, a factor of $\exp\{\hat{\theta}\hat{\Lambda}_k(\tau \wedge X_{ik\ell})\}$ appears to be missing in the definition of $\hat{\epsilon}_i$, on p. 392.

In view of the numerous difficulties encountered in trying to implement Glidden's test, it was ultimately decided to restrict attention to the corrected version of Shih's procedure. The power of the latter test is compared in Table2.6 with those of the tests based on S_n, T_n and S_{0n}. As might have been expected, differences in power between the four procedures are tenuous in cases of strong association. When the dependence is weak, however, the (corrected) Shih statistic turns out to be significantly more powerful than the other three, except when the alternative is Gumbel–Hougaard.

Table 2.6: Percentage of rejection of the null hypothesis of Clayton's copula for tests based on S_n, T_n, S_{0n}, and Shih's statistic at the 5% level when $n = 250$ for various copula alternatives, based on 10,000 replicates

Alternative		Estimated power			
Family	τ	S_n	T_n	S_{0n}	Shih
Clayton	0.20	4.6	5.5	4.9	4.2
	0.40	4.2	4.7	3.6	4.2
	0.60	4.6	4.2	5.0	4.9
	0.80	4.7	4.4	5.6	5.9
Frank	0.20	47.1	38.9	51.1	73.6
	0.40	97.1	88.4	97.4	99.9
	0.60	99.9	98.2	100.0	100.0
	0.80	100.0	96.2	100.0	100.0
Gumbel–Hougaard	0.20	68.4	57.6	80.6	8.8
	0.40	99.8	97.9	100.0	37.1
	0.60	100.0	100.0	100.0	78.8
	0.80	100.0	100.0	100.0	97.3
Fréchet	0.20	35.0	21.9	41.8	36.1
	0.40	81.7	60.2	89.2	85.1
	0.60	98.0	83.1	98.9	97.2
	0.80	98.3	86.3	99.4	98.3
Gaussian	0.20	33.3	23.7	38.7	52.3
	0.40	86.9	71.6	90.7	97.3
	0.60	99.1	92.1	99.8	100.0
	0.80	100.0	98.0	100.0	100.0
Plackett	0.20	49.5	35.1	49.7	70.5
	0.40	94.0	83.0	95.3	99.8
	0.60	99.5	94.8	99.8	100.0
	0.80	99.9	93.5	99.7	100.0

2.6 Illustrations

This section presents two illustrations of the proposed methodology to data sets originally considered by Frees & Valdez (1998) and by Cook & Johnson (1981, 1986).

2.6.1 Insurance data

Figure 2.1 displays the relation between the natural logarithms of an indemnity payment X_1 and an allocated loss adjustment expense X_2 (comprising lawyers' fees and claim investigation expenses, among others) for 1500 general liability claims. These data were used by Frees & Valdez (1998), Klugman & Parsa (1999) and Chen & Fan (2005), among others, to illustrate copula-model selection and fitting in an insurance context. In their analysis, Frees & Valdez (1998) ignored the censoring present in 34 claims in their visual procedure for the selection of an appropriate copula model, although they used the full sample in their formal estimation of the dependence parameter in the Clayton, Frank, and Gumbel–Hougaard copulas. (Although this is irrelevant here, they used generalized Pareto distributions for the marginals.)

For simplicity, the analysis presented in the sequel is limited to the 1466 uncensored claims. This restriction has little effect on the estimation of the dependence parameters, as evidenced by a comparison of the numerical estimates obtained by Frees & Valdez (1998) and Genest et al. (1998) with and without censoring, respectively. For the uncensored sample, the observed value of Kendall's tau is 0.3195, which is also the estimate of the dependence parameter θ in the Gumbel–Hougaard model. Frees & Valdez (1998) identified this model as providing the best fit of the three. Their judgment was based on a visual comparison of $K_n(t)$ and the parametric distribution functions $K(\theta_n, t)$ corresponding to the Clayton, Frank, and Gumbel–Hougaard dependence structures. The same conclusion was reached by Genest et al. (1998) and by Chen & Fan (2005) using more formal, pseudo-likelihood ratio based, procedures.

Figure 2.1: Scatter plot of the natural logarithms of the indemnity payment (LOSS) and the allocated loss adjustment expense (ALAE) for 1500 general liability claims

The nonparametric estimator $K_n(t)$ of $K(t)$ is shown in Figure 2.2, along with the parametric curves $K(\theta_n, t)$ corresponding to the Clayton, Frank, and Gumbel–Hougaard models, with θ_n estimated in each case through inversion of τ_n, the empirical version of Kendall's tau. The graph clearly suggests that the Gumbel–Hougaard copula is preferable. This conclusion is confirmed by formal tests based on S_n, T_n and Wang and Wells' statistic $S_{\xi n}$ with $\xi = 0$.

Figure 2.2: Nonparametric estimator $K_n(t)$ for the LOSS and ALAE data, along with three parametric estimators $K(\theta_n, t)$ corresponding to the Clayton, Frank, and Gumbel–Hougaard copula models, with θ_n estimated by inversion of the empirical version τ_n of Kendall's tau

The critical points and P-values reported in Table 2.7 were derived using $N = 10,000$ repetitions of the parametric bootstrap procedure described in section 4.2, which is based on C_n. While the P-values of S_n, T_n and S_{0n} lead to rejection of the Clayton and Frank dependence structures at the 5% level, they are larger than 0.8 for the Gumbel–Hougaard model.

Table 2.7: Results of the goodness-of-fit tests based on the statistics S_n, T_n and $S_{\xi n}$ with $\xi = 0$ for the data of LOSS and ALAE insurance data

Model	θ_n	S_n T_n S_{0n}	Critical value $c_{2n}(0.95)$	P-value (in %)
Clayton	0.939	2.330	0.135	0.0
		2.517	0.910	0.0
		1.892	0.126	0.0
Frank	3.143	0.244	0.123	0.0
		0.903	0.873	3.6
		0.330	0.128	0.0
Gumbel–Hougaard	0.319	0.027	0.117	88.8
		0.483	0.902	84.0
		0.051	0.127	90.2

Additional evidence in favor of the Gumbel–Hougaard extreme-value structure is supplied by Figure 2.3, on which the nonparametric estimator $K_n(t)$ is displayed, along with a global 95% confidence band for each of the three Archimedean models considered. Its limits are of the form

$$K(\theta_n, t) \pm \frac{1}{\sqrt{n}}\, c_{2n}(0.95), \qquad (2.13)$$

where $c_{2n}(0.95)$ is the 95% percentile of T_n under the null hypothesis, as reported in Table 2.7.

Figure 2.3: Nonparametric estimator $K_n(t)$ for the LOSS and ALAE data, along with global 95% confidence bands based on T_n for the three models considered: Clayton (top left panel), Frank (top right panel), and Gumbel–Hougaard (bottom panel)

2.6.2 Uranium exploration data

As a second illustration, the analysis of the uranium exploration data set originally considered by Cook & Johnson (1981, 1986) was revisited. These data consist of 655 chemical analysis from water samples collected from the Montrose quadrangle of western Colorado (USA). Concentrations were measured for the following elements: uranium (U), lithium (Li), cobalt (Co), potassium (K), cesium (Cs), scandium (Sc), and titanium (Ti).

Table 2.8 shows the values of the test statistics S_n, T_n and S_{0n} for selected pairs of variables, along with the corresponding P-values. The latter are based on $N = 10{,}000$ repetitions of the parametric bootstrap procedure.

The following observations can be drawn from Table 2.8 :

a) In the authors' experience, the three tests are generally in agreement, as for the pairs (U, Li) and (Co, Ti). Occasionally, they lead to different choices of models, as for the pairs (U, Co) and (Li, Sc).

b) The tests' P-values can sometimes differ markedly. This is often inconsequential, as in the case of the Gumbel–Hougaard copula for the pair (U, Li). In other occasions, however, the choice of statistic could make a difference between acceptance and rejection at a given level. At the 5% level, for example, Frank's model is acceptable for the pair (Co, Ti), both according to S_n and S_{0n}, but not under T_n. A similar phenomenon can be observed in the pair (Li, Ti), for which the four models considered would be accepted at the 15% level if T_n were used, but rejected by the other two statistics.

c) Although the statistics S_{0n} computed for different models have different distributions, it can be observed empirically that the model for which the statistic is smallest generally has the highest P-value. The pairs (U, Sc) and (Li, Ti) provide counterexamples which might or might not be due to sampling error. The same phenomenon could be observed for the statistics S_n and T_n; see, for example, pairs (U, Sc) and (Li, Ti).

2.7 Discussion

The goodness-of-fit procedures proposed herein will be consistent so long as $K(t) = P\{H(\mathbf{X}) \leq t\}$ assumes different functional forms under the hypothesized copula model and the true one. As already argued by Wang & Wells (2000) in the case of $S_{\xi n}$, such is the case for bivariate Archimedean copulas, a result which stems from the fact established by Genest & Rivest (1993) that the Archimedean generator ϕ is completely determined by K. This argument extends readily to S_n and T_n, or any other continuous func-

Table 2.8: Values taken by the goodness-of-fit statistics S_n, T_n and S_{0n} and associated P-values, for selected pairs in the uranium exploration data

Pair	Model	S_n	P-value (in %)	T_n	P-value (in %)	S_{0n}	P-value (in %)
(U, Li)	Ali–Mikhail–Haq	0.0880	22.4	0.6723	43.0	0.0690	33.0
	Clayton	0.3328	0.2	1.2329	0.5	0.2175	0.4
	Frank	**0.0538**	**52.5**	**0.5742**	**67.0**	**0.0448**	**74.2**
	Gumbel–Hougaard	0.1033	14.1	0.6727	42.5	0.1190	5.0
(U, Co)	Ali–Mikhail–Haq	**0.0836**	**24.7**	0.8328	13.7	**0.0679**	**33.4**
	Clayton	0.1057	20.6	**0.7730**	**28.3**	0.0793	24.7
	Frank	0.1099	10.7	0.8614	9.5	0.0862	12.6
	Gumbel-Hougaard	0.1448	6.6	0.9307	6.1	0.1351	4.2
(U, Sc)	Ali–Mikhail–Haq	0.2344	1.0	**1.2362**	0.2	**0.2077**	0.3
	Clayton	0.4042	0.2	1.3436	0.2	0.2934	0.1
	Frank	**0.2285**	**1.2**	1.2402	**0.3**	0.2140	**0.4**
	Gumbel–Hougaard	0.3203	0.1	1.4487	0.1	0.3470	0.0
(Li, Sc)	Ali–Mikhail–Haq	0.1347	6.4	0.9182	5.6	**0.0716**	**26.0**
	Clayton	**0.1120**	**16.6**	**0.8083**	**19.5**	0.0891	17.2
	Frank	0.1634	3.1	1.0443	1.6	0.0817	16.3
	Gumbel–Hougaard	0.2187	0.8	1.1089	0.9	0.1426	1.9
(Li, Ti)	Ali–Mikhail–Haq	0.1493	6.0	0.7456	28.1	0.1553	1.8
	Clayton	0.1382	10.8	0.7724	29.6	0.1578	3.4
	Frank	**0.1256**	12.0	**0.6463**	**53.0**	0.1401	**5.2**
	Gumbel–Hougaard	0.1327	**12.1**	0.6942	41.8	**0.1330**	4.9
(Co, Ti)	Ali–Mikhail–Haq	0.6294	0.0	1.5078	0.1	0.5916	0.0
	Clayton	0.6916	0.0	1.5009	0.0	0.5437	0.0
	Frank	**0.0731**	**23.0**	**0.9230**	**2.9**	**0.0539**	**48.5**
	Gumbel–Hougaard	0.2252	0.0	0.9899	1.4	0.2687	0.0

tional of the process $\mathbb{K}_n$. It may be conjectured that this characterization of Archimedean generators extends to arbitrary dimension $d \geq 2$. For, in the light of Equation (2.11) from Barbe et al. (1996), one can check easily that $y = \phi^{-1}$ is a solution of the differential equation

$$\sum_{i=0}^{d-1} \frac{(-1)^i}{i!} t^i y^{(i)}(t) - K\{\theta, y(t)\} = 0, \tag{2.14}$$

of order $d-1$ with boundary conditions $y(0) = 1$, and $t^i y^{(i)}(t) \to 0$ as $t \to \infty$ for all $i \in \{1, \dots, d-1\}$. It follows that if, given $K(\theta, \cdot)$, the solution of (2.14) is unique up to a scaling parameter, then the copula C_θ associated with $K(\theta, \cdot)$ is unique. In other words, the copula is uniquely determined by K whenever any two solutions y_1 and y_2 of the above equation satisfy $y_2(t) = y_1(\alpha t)$, for some $\alpha > 0$,

Nevertheless, there are circumstances when goodness-of-fit tests based on the process $\mathbb{K}_n$ will not be consistent. Suppose, for example, that C_1 and C_2 are two extreme-value copulas with the same value, τ, of Kendall's tau. In such a case, one would have $K(t) = t - (1 - \tau)t \log(t)$ for both models. Although the empirical process $\mathbb{K}_n$ may not have the same asymptotic distribution according as the data arise from C_1 or C_2, the limit would be a centered Gaussian process in both cases. Accordingly, the power of test statistics such as S_n and T_n, or indeed any other continuous functional of $\mathbb{K}_n$, could not approach one as $n \to \infty$.

One important advantage of the procedures proposed herein is that they are applicable to situations involving more than two variables. As an illustration, the goodness-of-fit of trivariate Ali–Mikhail–Haq, Clayton, Frank, and Gumbel–Hougaard copula models was checked on the triplet (Li, K, Ti).

Table 2.9 summarizes the results of the tests based on statistics S_n and T_n with $d = 3$. The P-values associated with the Frank and Gumbel–Hougaard dependence structures clearly lead to the rejection of those models. The Ali–Mikhail–Haq copula is also rejected at the 5% level by the Kolmogorov–Smirnov test. And indeed, for these three models, it may also be checked graphically (figures not provided) that the nonparametric estimator K_n lies in part outside the global 95% confidence band. There is, however, no evidence to conclude that formula

$$C_\theta(u_1, u_2, u_3) = \left(u_1^{-\theta} + u_2^{-\theta} + u_3^{-\theta} - 2\right)^{-1/\theta}, \quad \theta > 0 \tag{2.15}$$

should be discarded as a potential model for these data.

Table 2.9: Results of the goodness-of-fit tests based on the Cramér–von Mises and Kolmogorov–Smirnov statistics S_n and T_n for the trivariate data involving concentrations of lithium, potassium and titanium

Model	θ_n	S_n / T_n	critical value	P-value (in %)
Ali–Mikhail–Haq	0.242	1.106	1.370	14.8
		2.184	2.154	4.3
Clayton	0.122	0.264	0.837	50.8
		1.225	1.911	47.2
Frank	0.548	1.140	0.724	0.6
		2.193	1.702	0.5
Gumbel–Hougaard	0.055	1.347	0.656	0.0
		2.235	1.645	0.2

In subsequent work, it would be of interest to extend the set of copulas for which hypotheses I–IV are met. Because of their popularity in survival data analysis, where their mixture representation allows them to be viewed as a natural extension of Cox's proportional hazards model (Oakes, 2001), multi-parameter Archimedean models such as those considered by Joe (1997) and Genest et al. (1998) should probably be considered first. For the (Li, K, Ti) data considered just above, for example, it may well be that a multi-parameter copula model with different Clayton marginals for different pairs of variables (see for example Bandeen-Roche & Liang 1996) might be more appropriate than the somewhat restrictive (2.15).

CHAPITRE 3

ON THE JOINT ASYMPTOTIC BEHAVIOR OF TWO RANK-BASED ESTIMATORS OF THE ASSOCIATION PARAMETER IN THE GAMMA FRAILTY MODEL

Résumé

Dans le cadre de l'étude de simulation conduite au chapitre 2, il a été remarqué que l'expression pour la variance asymptotique donnée par Shih (1998) pour une statistique d'adéquation qu'elle a proposée est erronée. Le test de Shih est basé sur les estimateurs de rangs proposés par Clayton (1978) et par Oakes (1982) pour le paramètre d'association du modèle de fragilité gamma bivarié. Dans cet article, la loi asymptotique jointe de ces estimateurs est considérée, au moyen d'une approche différente de celle utilisée par Oakes (1982, 1986). Ceci mène à une correction de la formule donnée par Shih (1998) pour la covariance limite entre ces deux estimateurs.

Abstract

Rank-based estimators were proposed by Clayton (1978) and Oakes (1982) for the association parameter in the bivariate gamma frailty model. The joint asymptotic behavior of these estimators is considered here, following a different approach from that used by Oakes (1982, 1986). This leads to a correction of the formula given by Shih (1998) for the limiting covariance between the two estimators.

3.1 Introduction

A bivariate distribution function C with uniform margins is said to belong to the class $\mathcal{C}$ of (positively dependent) Clayton copulas if it can be written in the form

$$C(u,v) = \left(u^{-a} + v^{-a} - 1\right)^{-1/a}, \quad u,v \in (0,1)$$

for some $a > 0$. This family of dependence functions, for which the alternative parameterizations $a = \theta - 1 = 1/\eta$ are sometimes convenient, is popular in survival analysis, where it is generally referred to as the gamma frailty model.

This note reconsiders the joint asymptotic behavior of two rank-based estimators of the association parameter θ, based on a random sample (X_1, Y_1), ..., (X_n, Y_n) from a distribution H with continuous margins whose underlying copula belongs to $\mathcal{C}$. The first estimator, introduced by Oakes (1982), is based on an inversion of Kendall's tau statistic. The second estimator, due to Clayton (1978), stems from a pseudo-likelihood approach. Clayton & Cuzick (1985) later showed that it could be expressed as a weighted form of Oakes' concordance-based estimator.

The asymptotic behavior of the two estimators was originally considered by Oakes (1982, 1986). It is revisited here using a different approach which may be of independent interest. As briefly shown in Sections 3.2 and 3.3, the new approach yields the same limiting expressions as reported by Oakes for the variance of the rank-based estimators of $\log(\theta)$ in Clayton's model. In Section 4, however, an expression obtained for the asymptotic covariance between the two estimators is seen to differ substantially from that reported by Shih (1998). The error in the latter can lead to a negative limiting variance for the goodness-of-fit statistic she suggests for judging the adequacy of Clayton's model. The point has apparently gone unnoticed so far but is worth rectifying, because Shih's test is used regularly in practice; see, e.g., Fine et al. (2001), He & Lawless (2003) or Wang (2003).

Before proceeding, note that since Clayton's and Oakes' estimators are rank-based, their limiting distribution does not depend on the margins F and G of the X_i and the Y_i, but only on the copula C joining $U_i = F(X_i)$ and $V_i = G(Y_i)$. Accordingly, one may take $H = C$ without loss of generality in the sequel.

3.2 Oakes' concordance estimator

It is well known that in Clayton's model, the population value

$$\tau = 4\mathrm{E}\{H(X,Y)\} - 1$$

of Kendall's tau is related to the dependence parameter through the identities

$$\tau = \frac{\theta - 1}{\theta + 1} = \frac{a}{a+2} = \frac{1}{1+2\eta}\,.$$

In particular, $\theta = (1+\tau)/(1-\tau)$; see, e.g., Nelsen (1999), p. 130.

Now if

$$\Delta_{ij} = \mathbf{1}(U_i < U_j, V_i < V_j) + \mathbf{1}(U_j < U_i, V_j < V_i)$$

for any distinct $i, j \in \{1, \ldots, n\}$, the standard empirical version of τ is

$$\tau_n = -1 + 2 \sum_{i<j} \Delta_{ij} \Big/ \binom{n}{2}.$$

This led Oakes (1982) to propose $\tilde{\theta}_n = (1 + \tau_n)/(1 - \tau_n)$ as a rank-based estimate of θ. This amounts to estimating θ by the ratio of the number of concordant pairs to the number of discordant pairs in the sample.

The asymptotic behavior of the U-statistic τ_n is well known to be normal. A standard application of Hájek's projection method (see Hájek & Šidák 1967) yields

$$\sqrt{n}\,(\hat{\tau}_n - \tau) = \frac{4}{\sqrt{n}} \sum_{i=1}^{n} \left\{ C(U_i, V_i) + \bar{C}(U_i, V_i) - \frac{\theta}{\theta + 1} \right\} + o_P(1),$$

where $\bar{C}(u, v) = 1 - u - v + C(u, v)$ is the survival function associated with the Clayton copula $C \in \mathcal{C}$ with dependence parameter θ. Hence, by Slutsky's theorem,

$$\sqrt{n} \left\{ \log(\tilde{\theta}_n) - \log(\theta) \right\}$$
$$= \frac{2(1+\theta)^2}{\theta\sqrt{n}} \sum_{i=1}^{n} \left\{ C(U_i, V_i) + \bar{C}(U_i, V_i) - \frac{\theta}{\theta + 1} \right\} + o_P(1).$$

Invoking Hoeffding's identity (see, e.g., Nelsen 1999, p. 154), one gets also

$$\mathrm{E}(UV) = \int_0^1 \int_0^1 C(u, v)\, \mathrm{d}v\, \mathrm{d}u = \eta^2 L(\eta),$$

where

$$L(\eta) = \sum_{k=0}^{\infty} \frac{\Gamma^2(\eta)}{\Gamma(\eta)} \frac{k!\Gamma(\eta + k)}{\Gamma(2\eta + k + 1)} = \frac{1}{4\eta^2}\,\text{hypergeom}([1, 1, \eta], [2\eta + 1, 2\eta + 1], 1).$$

In particular, note that $\eta^2 L(\eta) \to 1/4$ as $\eta \to \infty$, corresponding to independence.

Consequently, the limiting distribution of $\sqrt{n}\,\{\log(\tilde{\theta}_n) - \log(\theta)\}$ is Gaussian, with mean zero and variance

$$V(\eta) = \frac{8(2\eta+1)^4}{(\eta+1)^2}\,L(\eta) - \frac{4}{3}\frac{(2\eta+1)^2(17\eta^3+27\eta^2+14\eta+2)}{\eta^2(\eta+1)^2(3\eta+1)}\,,$$

as reported in equation (5) of Shih (1998). However, the series representation of L given in her article is incorrect.

3.3 Clayton's weighted concordance estimator

As shown by Clayton & Cuzick (1985), the pseudo-likelihood estimator of θ originally proposed by Clayton (1978) can be written in the form

$$\hat{\theta}_n = \frac{\sum_{i<j} \Delta_{ij}/R_{ij}}{\sum_{i<j}(1-\Delta_{ij})/R_{ij}} = \frac{A_n}{B_n}\,,$$

where $R_{ij} = nC_n\left\{\max(U_i, U_j), \max(V_i, V_j)\right\}$ and

$$C_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(U_i \le u, V_i \le v).$$

Clayton's estimator thus differs from Oakes' in that it assigns a higher weight to small observations than to large ones.

The limiting behavior of A_n and B_n is considered in Sections 3.3.1 and 3.3.2, respectively. The asymptotic distribution of $\log(\hat{\theta}_n)$ is then determined in Section 3.3.3.

3.3.1 Asymptotic behavior of A_n

It is argued here that $\sqrt{n}\,(1 - A_n/n) \to 0$ in probability as $n \to \infty$, so that the limiting behavior of $\hat{\theta}_n$ is the same as that of n/B_n. Writing

$$\frac{\Delta_{ij}}{R_{ij}} = \frac{\mathbf{1}(U_i < U_j, V_i < V_j)}{nC_n(U_j, V_j)} + \frac{\mathbf{1}(U_i > U_j, V_i > V_j)}{nC_n(U_i, V_i)},$$

one sees that

$$A_n = \sum_{i \neq j} \frac{\mathbf{1}(U_i < U_j, V_i < V_j)}{nC_n(U_j, V_j)} = \sum_{j=1}^{n} \frac{nC_n(U_j, V_j) - 1}{nC_n(U_j, V_j)} = n - \frac{1}{n} \sum_{j=1}^{n} \frac{1}{C_n(U_j, Y_j)}.$$

Hence

$$1 - \frac{A_n}{n} = \frac{1}{n^2} \sum_{j=1}^{n} \frac{1}{C_n(U_j, V_j)} \geq 0. \tag{3.1}$$

The conclusion thus follows if it can be shown that $\sqrt{n}\,\mathrm{E}(1 - A_n/n) \to 0$. To this end, introduce $W = C(U_1, V_1)$, whose density equals $(a+1)\,(1 - w^a)\,/a$ for $w \in [0, 1]$; see, e.g., Genest & Rivest (2001). Then

$$\begin{aligned}
\mathrm{E}\left(1 - \frac{A_n}{n}\right) &= \mathrm{E}\left\{\frac{1}{nC_n(U_1, V_1)}\right\} \\
&= \sum_{k=0}^{n-1} \binom{n-1}{k} \frac{1}{k+1}\,\mathrm{E}\left\{W^k (1 - W)^{n-1-k}\right\} \\
&= \mathrm{E}\left\{\frac{1 - (1 - W)^n}{nW}\right\} \\
&\leq \left(\frac{a+1}{an}\right) \left\{\int_0^{1/n} \frac{1 - (1 - w)^n}{nw}\,\mathrm{d}w + \int_{1/n}^{1} \frac{1}{nw}\,\mathrm{d}w\right\} \\
&\leq \left(\frac{a+1}{an}\right) \left\{\int_0^{1} \frac{1 - e^{-2u}}{u}\,\mathrm{d}u + \log(n)\right\} \\
&= O\left\{\frac{\log(n)}{n}\right\}.
\end{aligned}$$

3.3.2 Asymptotic behavior of n/B_n

It follows from the definition of B_n that

$$\frac{B_n}{n} = \frac{1}{n^2} \sum_{i<j} \frac{1-\Delta_{ij}}{C_n\{\max(U_i,U_j),\max(V_i,V_j)\}} + o_P(1/\sqrt{n}\,).$$

Given the tightness of $\sqrt{n}\,(C_n - C)$, one may write

$$\frac{B_n}{n} = D_n - E_n + o_P\left(1/\sqrt{n}\right),$$

where

$$D_n = \frac{1}{n^2} \sum_{i \neq j} \frac{\mathbf{1}(U_i < U_j, V_i > V_j)}{C(U_j, V_i)}$$

and, in view of Corollary 2.5 of Ghoudi & Rémillard (2004),

$$E_n \;=\; \int \frac{\mathbf{1}(u_1 < u_2, v_1 > v_2)\,\{C_n(u_2,v_1) - C(u_2,v_1)\}}{C^2(u_2,v_1)}\,\mathrm{d}C(u_2,v_2)\,\mathrm{d}C(u_1,v_1).$$

A standard argument from the theory of U-statistics (e.g., Theorem 1 of Lee 1990, p. 76) then yields

$$D_n = \frac{1}{n} \sum_{i=1}^{n} \left\{ \log(U_i V_i) + \frac{2}{a}\,\log\left(U_i^{-a} + V_i^{-a} - 1\right) - \frac{2}{a+1} \right\} + o_P\left(1/\sqrt{n}\right).$$

Now because of the specific form of $C \in \mathcal{C}$, one can check easily that

$$\int \frac{\mathbf{1}(u_1 < u_2, v_1 > v_2)\mathbf{1}(U_i \le u_2, V_i \le v_1)}{C^2(u_2,v_1)}\,\mathrm{d}C(u_2,v_2)\,\mathrm{d}C(u_1,v_1)$$

$$= -\frac{1}{a}\log(U_i V_i) - \frac{1}{a^2}\log\left(U_i^{-a} + V_i^{-a} - 1\right)$$

and

$$\int \frac{\mathbf{1}(u_1 < u_2, v_1 > v_2)}{C(u_2, v_1)} \, \mathrm{d}C(u_2, v_2) \, \mathrm{d}C(u_1, v_1) = \frac{1}{a+1}.$$

As a result, E_n may be expressed as a sum of mutually independent and identically distributed, centered terms, viz.

$$E_n = -\frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{1}{a} \log(U_i V_i) + \frac{1}{a^2} \log\left(U_i^{-a} + V_i^{-a} - 1\right) + \frac{1}{a+1} \right\}.$$

3.3.3 Asymptotic behavior of $\log(\hat{\theta}_n)$

In view of the previous developments, it is clear that $\sqrt{n}\,(1/\hat{\theta}_n - 1/\theta)$ and

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{a+1}{a} \log(U_i V_i) + \frac{2a+1}{a^2} \log\left(U_i^{-a} + V_i^{-a} - 1\right) - \frac{1}{a+1} \right\}$$

have the same asymptotic distribution. An application of Slutsky's theorem then implies that $\sqrt{n}\{\log(\hat{\theta}_n) - \log(\theta)\}$ and

$$-\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \frac{(\eta+1)^2}{\eta} \log(U_i V_i) + (\eta+1)(\eta+2) \log\left(U_i^{-1/\eta} + V_i^{-1/\eta} - 1\right) - 1 \right\}$$

also have the same limiting behavior. Consequently, the large-sample distribution of $\sqrt{n}\{\log(\hat{\theta}_n) - \log(\theta)\}$ is Gaussian, with mean zero and variance

$$\begin{aligned} V_w(\eta) &= 2\eta^2 + 6\eta + 5 - 4(\eta+1)^4 \sum_{k=0}^{\infty} \frac{(-1)^k}{(\eta+k+1)^2} && (3.2) \\ &= 2\eta^2 + 6\eta + 5 - (\eta+1)^4 \left\{\Psi'(1+\eta/2) - \Psi'(1/2+\eta/2)\right\}, && (3.3) \end{aligned}$$

where Ψ' denotes the trigamma function, i.e., the derivative of the digamma function Ψ. Formula (3.3) is that which is reported by Oakes (1986). In the authors' experience, however, the form (3.2) is numerically more stable.

3.4 Asymptotic covariance between $\log(\tilde{\theta}_n)$ and $\log(\hat{\theta}_n)$

The developments presented thus far would be pointless, if their only purpose were to confirm the variance calculations already reported by Oakes (1982, 1986). As will be seen presently, however, they also allow for a straightforward calculation of the asymptotic covariance between $\log(\tilde{\theta}_n)$ and $\log(\hat{\theta}_n)$. The latter is given by

$$H(\eta) = -\frac{8\eta^3 + 19\eta^2 + 15\eta + 3}{\eta^2} + 4(\eta+1)(2\eta+1)^2 J(\eta), \tag{3.4}$$

where

$$\begin{aligned} J(\eta) &= \frac{1}{\eta^2}\,\mathrm{E}\left\{U \log(V)\right\} + \frac{1}{\eta^2} \\ &= \sum_{k=0}^{\infty} \frac{\Gamma(2\eta)k!}{(\eta+k)\Gamma(2\eta+k+1)} \\ &= \frac{1}{2\eta^2}\,\text{hypergeom}\left([1,1,\eta],[\eta+1,2\eta+1],1\right). \end{aligned}$$

This formula for $H(\eta)$ differs markedly from that given by Shih (1998), viz.

$$H^\star(\eta) = -4\eta - \frac{40\eta^3 + 49\eta^2 + 21\eta + 3}{(2\eta+1)\eta^2} + 4(\eta+1)(2\eta+1)^2 J(\eta).$$

The latter is erroneous. In particular, note that

$$H^\star(\eta) - H(\eta) = \frac{2\eta(4\eta+1)}{2\eta+1} \to \infty$$

as $\eta \to \infty$. Using $H(\eta)$, it may be checked, e.g., that the covariance between $\log(\tilde{\theta}_n)$ and $\log(\hat{\theta}_n)$ tends to $16\text{cov}^2\{U, \log(U)\} = 1$ as $\eta \to \infty$.

To establish (3.4), start from the definition of $H(\eta)$, viz.

$$-2\frac{(2\eta+1)^2}{\eta(\eta+1)}\,\mathrm{E}\left[\left\{2C(U,V)-U-V+1-\frac{\eta+1}{2\eta+1}\right\}\right.$$

$$\left.\times\left\{\frac{(\eta+1)^2}{\eta}\log(UV)+(\eta+1)(\eta+2)\log\left(U^{-1/\eta}+V^{-1/\eta}-1\right)-1\right\}\right]$$

$$\begin{aligned} = \quad & -8\,\frac{(2\eta+1)^2(\eta+1)}{\eta^2}\,\mathrm{cov}\{C(U,V),\log(U)\} \\ & + 2\,\frac{(2\eta+1)^2(\eta+1)}{\eta^2}\,\mathrm{cov}\{U+V,\log(UV)\} \\ & - 4\,\frac{(2\eta+1)^2(\eta+2)}{\eta}\,\mathrm{cov}\left\{C(U,V),\log\left(U^{-1/\eta}+V^{-1/\eta}-1\right)\right\} \\ & + 4\,\frac{(2\eta+1)^2(\eta+2)}{\eta}\,\mathrm{cov}\left\{U,\log\left(U^{-1/\eta}+V^{-1/\eta}-1\right)\right\}. \end{aligned}$$

Each of these summands can be evaluated relatively easily. To compute the first term, the change of variables $x=u^{-1/\eta}$, $y=v^{-1/\eta}$ may be used to write

$$\mathrm{E}\{\log(U)C(U,V)\} = -\eta^2(\eta+1)\int_1^\infty\int_1^\infty \log(x)(x+y-1)^{-2\eta-2}\,\mathrm{d}y\,\mathrm{d}x,$$

which, upon integration with respect to y, yields

$$-\frac{\eta^2(\eta+1)}{2\eta+1}\int_1^\infty x^{-2\eta-1}\log(x)\,\mathrm{d}x = -\frac{\eta+1}{8\eta+4}\,.$$

Therefore,

$$\mathrm{cov}\{\log(U),C(U,V)\} = \frac{\eta+1}{8\eta+4}\,,$$

since $\mathrm{E}\{\log(U)\}=-1$ and

$$\mathrm{E}\{C(U,V)\} = \mathrm{E}(W) = \frac{\tau+1}{4} = \frac{\eta+1}{4\eta+2}\,.$$

For the second summand, let $x = u^{-1/\eta} - 1$ and $y = v^{-1/\eta} - 1$ in $\mathrm{E}\{U \log(V)\} = \mathrm{E}\{V \log(U)\}$. Since $\mathrm{cov}\{U, \log(U)\} = \mathrm{cov}\{V, \log(V)\} = 1/4$, one gets

$$\mathrm{cov}\{U + V, \log(UV)\} = \frac{3}{2} + 2\,\mathrm{E}\{U \log(V)\} = -\frac{1}{2} + 2\eta^2 J(\eta).$$

To treat the third term, one can set once more $x = u^{-1/\eta}$ and $y = v^{-1/\eta}$ to get

$$\begin{aligned} &\mathrm{E}\left\{C(U,V)\log\left(U^{-1/\eta} + V^{-1/\eta} - 1\right)\right\} \\ &\qquad = \eta(\eta+1)\int_1^\infty \int_1^\infty (x + y - 1)^{-2\eta-2} \log(x + y - 1)\,\mathrm{d}y\,\mathrm{d}x. \end{aligned}$$

Letting $s = x$ and $t = x + y - 1$, one may then compute this expectation as follows:

$$\begin{aligned} \eta(\eta+1)\int_1^\infty \int_1^t t^{-2\eta-2}\log(t)\,\mathrm{d}s\,\mathrm{d}t &= \eta(\eta+1)\int_1^\infty (t-1)t^{-2\eta-2}\log(t)\,\mathrm{d}t \\ &= \frac{(\eta+1)(4\eta+1)}{4\eta(4\eta^2+4\eta+1)}. \end{aligned}$$

Consequently,

$$\mathrm{cov}\left\{C(U,V),\ \log\left(U^{-1/\eta} + V^{-1/\eta} - 1\right)\right\} = -\frac{4\eta^2 + 3\eta + 1}{4\eta(4\eta^2 + 4\eta + 1)}\,.$$

The covariance in the fourth summand can be handled with the same succession of changes of variables. One finds

$$\mathrm{cov}\left\{U,\ \log\left(U^{-1/\eta} + V^{-1/\eta} - 1\right)\right\} = -\frac{4\eta^2 + 3\eta + 1}{4\eta(4\eta^2 + 4\eta + 1)} = -\frac{1}{4\eta}\,.$$

Formula (3.4) obtains upon gathering terms.

3.5 Closing comments

In her paper, Shih (1998) proposes the statistic $T_n = \sqrt{n}\log(\tilde{\theta}_n/\hat{\theta}_n)$ for testing the adequacy of Clayton's gamma frailty model. The asymptotic distribution of T_n is normal, with mean zero and variance

$$W(\eta) = V(\eta) + V_w(\eta) - 2H(\eta).$$

Upon substitution, one finds

$$\begin{aligned} W(\eta) & \\ = \; & \frac{18\eta^7 + 240\eta^6 + 3001\eta^5 + 8281\eta^4 + 9449\eta^3 + 5171\eta^2 + 1352\eta + 136}{3\eta^2(\eta+1)^2(3\eta+1)} \\ & + \frac{8(2\eta+1)^4}{(\eta+1)^2} L(\eta) - (\eta+1)^4 \left\{ \Psi'(1+\eta/2) - \Psi'(1/2+\eta/2) \right\} \\ & - 8(\eta+1)(2\eta+1)^2 J(\eta). \end{aligned}$$

This implies, e.g., that $W(\eta) \to 7/9$ as $\eta \to \infty$. By opposition, the use of $H^\star(\eta)$ instead of $H(\eta)$ would make $W(\eta)$ negative for large values of η and close to $-\infty$ in the neighborhood of independence.

Fortunately, the above correction does not affect the conclusions derived by Fine et al. (2001), He & Lawless (2003) or Wang (2003) in their applications of Shih's goodness-of-fit test.

CHAPITRE 4

TESTS D'ADÉQUATION BASÉS SUR LE PROCESSUS DE COPULE EMPIRIQUE

4.1 Introduction

L'étude de simulation conduite au chapitre 2 a montré la bonne performance des procédures d'adéquation basées sur $\mathbb{K}_n$ sous plusieurs modèles de contre-hypothèses exprimés à l'aide de copules. Le comportement de ces tests est également appréciable pour des contre-hypothèses provenant d'une version multivariée de la loi de Student, tel qu'on peut le constater au tableau 4.1. À l'instar des résultats observés au chapitre 2, la puissance du test d'adéquation fondé sur S_n semble supérieure à celle du test basé sur T_n.

Rappelons qu'un vecteur aléatoire T est distribué selon une loi de Student d-variée à ν degrés de liberté, de moyenne μ et de matrice de variance-covariance $\nu\Sigma/(\nu-2)$ si

$$T \stackrel{d}{=} \mu + \frac{\sqrt{\nu}}{\sqrt{S}}\, Y,$$

Table 4.1: Pourcentage de rejet de trois hypothèses nulles différentes avec S_n et T_n au niveau 5% et $n = 250$ pour des contre-hypothèses de Student bivariée à ν degrés de liberté, basé sur 1000 répétitions

Loi de Student		Modèle sous l'hypothèse nulle					
Degré de liberté	Valeur de ρ	Clayton		Frank		Gumbel–Hougaard	
		S_n	T_n	S_n	T_n	S_n	T_n
$\nu = 2$	0.2	20.9	14.3	47.5	41.7	41.1	41.1
	0.4	55.5	37.9	61.4	49.3	56.4	50.6
	0.6	88.2	73.0	73.3	59.5	67.6	57.2
	0.8	99.1	93.8	87.7	69.1	62.9	51.8
$\nu = 3$	0.2	14.7	11.3	31.0	28.6	30.7	29.9
	0.4	51.8	36.5	44.5	35.8	48.8	40.0
	0.6	86.4	71.5	58.2	46.3	61.5	48.8
	0.8	98.7	92.8	78.1	59.5	58.9	46.4
$\nu = 5$	0.2	13.3	9.9	19.2	18.4	23.4	21.5
	0.4	51.0	36.0	30.2	25.1	42.3	33.2
	0.6	86.5	71.2	44.8	35.6	56.6	42.5
	0.8	98.7	92.7	67.9	49.9	56.7	42.8
$\nu = 7$	0.2	12.1	9.4	15.7	14.7	21.7	18.4
	0.4	51.2	36.6	24.7	20.5	40.4	30.4
	0.6	86.2	71.2	39.2	32.1	56.1	41.9
	0.8	98.7	92.0	63.8	46.1	55.0	41.9
$\nu = 10$	0.2	12.5	9.8	12.8	12.5	19.7	16.9
	0.4	51.8	36.7	21.0	18.9	38.0	29.7
	0.6	87.3	70.9	34.7	27.5	54.6	38.8
	0.8	98.8	91.4	59.1	42.5	55.2	40.9
$\nu = 15$	0.2	13.1	10.1	10.8	10.7	18.1	15.6
	0.4	51.2	37.4	18.8	16.2	38.8	28.7
	0.6	86.7	71.5	30.7	25.2	54.3	39.2
	0.8	98.9	91.8	57.5	40.1	55.4	40.7

où la variable Y est distribuée selon une loi $N_d(0, \Sigma)$, indépendamment de S, qui est une variable khi-deux à ν degrés de liberté. La copule associée à T se déduit directement du théorème 1.1 dû à Sklar (1959).

Cependant, malgré la puissance élevée des méthodes d'adéquation basées sur S_n et T_n sous plusieurs variétés de contre-hypothèses, quelques critiques concernant les tests étudiés au chapitre 2 peuvent être soulevées:

(a) Puisque la transformation intégrale de probabilité est une projection unidimensionnelle d'une fonction de dépendance à d dimensions, les tests développés au chapitre 2 ne seront pas toujours convergents;

(b) Il se pourrait que les hypothèses qui assurent la convergence du processus d'adéquation ne soient pas satisfaites pour certaines classes importantes de copules;

(c) Pour certains modèles, la transformation intégrale de probabilité n'admet pas toujours de forme explicite.

En regard du point (a), on s'attend à ce que l'efficacité de S_n et de T_n décroisse à mesure que la dimension du vecteur augmente. Également, comme cas particulier de la remarque (c), on retrouve la copule normale, tirée du modèle multivarié classique par une application du théorème de Sklar.

Pour pallier ces limitations, la section suivante procure les outils théoriques permettant de développer des tests d'ajustement basés sur une estimation directe et complètement non paramétrique de la copule sous-jacente à une population. Des tests pouvant rivaliser avec S_n et T_n, de même qu'une procédure d'adéquation applicable à la copule normale, pourront alors être envisagés.

4.2 Processus de copule empirique pour l'adéquation

Pour répondre aux difficultés soulevées dans l'introduction, il est proposé ici d'appuyer une procédure de test sur une estimation non paramétrique C_n de C. Spécifiquement, si $(X_{11},\ldots,X_{1d}),\ldots,(X_{n1},\ldots,X_{nd})$ est un échantillon aléatoire d'une population d-variée, on définit la copule empirique par

$$C_n(u) = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d}\mathbf{1}\{F_{nj}(X_{ij}) \leq u_j\}, \tag{4.1}$$

où

$$F_{nj}(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\left(X_{ij} \leq x\right), \quad 1 \leq j \leq d$$

sont les fonctions de répartition empiriques marginales. Il est facile de voir que l'estimateur C_n est entièrement fondé sur les rangs des observations, puisque $nF_{nj}(X_{ij}) = R_{ij}$ pour tous $i \in \{1,\ldots,n\}$ et $j \in \{1,\ldots d\}$.

Soit maintenant une famille paramétrique $\{C_\theta;\, \theta \in \mathcal{O}\}$ de copules , où $\mathcal{O} \subseteq \mathbb{R}^m$. Afin d'obtenir la loi asymptotique d'un processus d'adéquation basé sur C_n, les trois hypothèses suivantes seront nécessaires.

Hypothèse 1. *Pour tout $\theta \in \mathcal{O}$, les dérivées partielles $\partial C_\theta(u)/\partial u_j$, $1 \leq j \leq d$ existent et sont continues sur $[0,1]^d$.*

Soit θ_n, un estimateur du paramètre inconnu θ.

Hypothèse 2. *La suite* $\Theta_n = \sqrt{n}\,(\theta_n - \theta)$ *converge en loi vers* Θ.

Hypothèse 3. *Pour tout* $\theta \in \mathcal{O}$,

$$\sup_{\|\theta^\star - \theta\| < \varepsilon} \sup_{u \in [0,1]^d} \left| \dot{C}_{\theta^\star}(u) - \dot{C}_\theta(u) \right| \longrightarrow 0$$

lorsque $\varepsilon \to 0$, *où*

$$\dot{C}_\theta = \nabla_\theta C_\theta = \left(\frac{\partial}{\partial \theta_1} C_\theta, \ldots, \frac{\partial}{\partial \theta_m} C_\theta \right)^\top,$$

est le gradient de C_θ *par rapport à* θ.

L'hypothèse 1 assure la convergence faible du processus empirique $\mathbb{C}_{n,\theta} = \sqrt{n}\,(C_n - C_\theta)$ vers une limite gaussienne $\mathbb{C}_\theta$. Ce résultat, dû à Gänssler & Stute (1987), a récemment été reconsidéré par Ghoudi & Rémillard (2004) et par Fermanian et coll. (2004).

Maintenant, pour tester l'adéquation à une famille paramétrique, on considère le processus d'adéquation $\mathbb{C}_n = \sqrt{n}\,(C_n - C_{\theta_n})$. Les hypothèses 2 et 3 permettent de caractériser la limite de ce processus.

Proposition 4.1. *Supposons que les hypothèses I–III sont satisfaites pour une famille* $\{C_\theta,\ \theta \in \mathcal{O}\}$ *de copules. Alors le processus empirique* $\mathbb{C}_n = \sqrt{n}\,(C_n - C_{\theta_n})$ *converge dans l'espace* $\mathcal{C}(\mathbb{R}^d)$ *vers un processus gaussien et centré* $\mathbb{C} = \mathbb{C}_\theta - \Theta \dot{C}_\theta$.

Démonstration. Considérons la décomposition $\mathbb{C}_n = \mathbb{C}_{n,\theta} - B_{n,\theta}$, où

$$\mathbb{C}_{n,\theta} = \sqrt{n}(C_n - C_\theta) \quad \text{et} \quad B_{n,\theta} = \sqrt{n}(C_{\theta_n} - C_\theta).$$

Dans un premier temps, puisque l'hypothèse 1 est satisfaite, le Théorème 4.1 implique que $\mathbb{C}_{n,\theta}$ converge vers $\mathbb{C}_\theta$.

Dans un deuxième temps, il sera démontré que $B_{n,\theta}$ peut devenir arbitrairement près de $\Theta_n \dot{C}_\theta$, en autant que n soit suffisamment grand. De l'hypothèse 2, Θ_n converge en loi vers Θ, ce qui implique que la suite Θ_n est tendue. Ainsi, pour un $\delta > 0$ arbitraire, il existe un $M = M_\delta$ telle que $\mathrm{P}\left(\|\Theta_n\| > M\right) < \delta$. Donc,

$$
\begin{aligned}
&\mathrm{P}\left\{\sup_{u\in[0,1]^d}\left|B_{n,\theta}(u) - \Theta_n \dot{C}_\theta(u)\right| > \gamma\right\} \\
&\quad \leq \mathrm{P}\left\{\sup_{u\in[0,1]^d}\left|B_{n,\theta}(u) - \Theta_n \dot{C}_\theta(u)\right| > \gamma,\ \|\Theta_n\| \leq M\right\} + \mathrm{P}\left(\|\Theta_n\| > M\right) \\
&\quad < \mathrm{P}\left\{\sup_{u\in[0,1]^d}\left|B_{n,\theta}(u) - \Theta_n \dot{C}_\theta(u)\right| > \gamma,\ \|\Theta_n\| \leq M\right\} + \delta.
\end{aligned}
$$

On déduit alors du théorème de la valeur moyenne que $B_{n,\theta} = \Theta_n \dot{C}_{\theta_n^\star}$, où $\theta_n^\star = \theta + \varepsilon n^{-1/2}\Theta_n$, $0 < \varepsilon < 1$, et en employant l'hypothèse 3,

$$
\begin{aligned}
&\lim_{n\to\infty} \mathrm{P}\left\{\sup_{u\in[0,1]^d}\left|B_{n,\theta}(u) - \Theta_n \dot{C}_\theta(u)\right| > \gamma,\ \|\Theta_n\| \leq M\right\} \\
&\qquad = \lim_{n\to\infty} \mathrm{P}\left\{\|\Theta_n\| \sup_{u\in[0,1]^d}\left|\dot{C}_{\theta_{n^\star}}(u) - \dot{C}_\theta(u)\right| > \gamma,\ \|\Theta_n\| \leq M\right\} \\
&\qquad \leq \lim_{n\to\infty} \mathrm{P}\left\{\sup_{\|\theta_n^\star - \theta\| \leq n^{-1/2}M}\ \sup_{u\in[0,1]^d}\left|\dot{C}_{\theta_{n^\star}}(u) - \dot{C}_\theta(u)\right| > \frac{\gamma}{M}\right\} \\
&\qquad = 0.
\end{aligned}
$$

Par conséquent, la limite du processus $B_{n,\theta}$ est la même que celle de $\Theta_n \dot{C}_\theta$, à savoir $\Theta \dot{C}_\theta$, ce qui complète la démonstration. ♢

4.3 Application aux copules archimédiennes

Les hypothèses 1 et 3, qui concernent la forme analytique des membres d'une famille de copules, sont satisfaites pour plusieurs familles de copules archimédiennes. Rappelons qu'une copule C appartenant à cette classe s'exprime sous la forme $C(u) = \phi^{-1}\{\phi(u_1) + \cdots + \phi(u_d)\}$, où le générateur $\phi : (0,1] \to [0,\infty)$ est une bijection qui satisfait que

$$\phi(1) = 0 \quad \text{et} \quad (-1)^i \frac{\mathrm{d}^i}{\mathrm{d}x^i}\,\phi^{-1}(x) > 0, \quad 1 \leq i \leq d. \tag{4.2}$$

Considérons maintenant une famille $\{C_\theta,\ \theta \in \mathcal{O}\}$ de copules archimédiennes, dont chaque membre est caractérisé par un générateur ϕ_θ qui vérifie (4.2), et définissons

$$\dot{\phi}_\theta = \nabla_\theta \phi_\theta = \left(\frac{\partial}{\partial \theta_1}\,\phi_\theta, \ldots, \frac{\partial}{\partial \theta_m}\,\phi_\theta \right)^\top,$$

le gradient of ϕ_θ par rapport à θ. La proposition suivante permet de représenter $\dot{C}_\theta$ uniquement en fonction de $\dot{\phi}_\theta$.

Proposition 4.2. *Si chaque élément de la famille $\{C_\theta;\, \theta \in \Theta\}$ est une copule archimédienne dont le générateur est ϕ_θ, alors pour chaque $\theta \in \mathcal{O}$, on a*

$$\dot{C}_\theta(u) = \frac{\dot{\phi}_\theta(u_1) + \cdots + \dot{\phi}_\theta(u_d) - \dot{\phi}_\theta\{C_\theta(u)\}}{\phi'_\theta\{C_\theta(u)\}}.$$

Démonstration. Le résultat s'obtient aisément en écrivant

$$\phi_\theta\{C_\theta(u)\} = \phi_\theta(u_1) + \cdots + \phi_\theta(u_d)$$

et en appliquant la règle de dérivation en chaîne. ◊

Exemple 4.1. *Le générateur de la copule de Clayton est défini par*

$$\phi_\theta(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right), \quad \theta \geq 0.$$

Il en découle que

$$\phi'_\theta(t) = -t^{-\theta-1} \quad et \quad \dot{\phi}_\theta(t) = \frac{1 - t^{-\theta} - \theta t^{-\theta}\log t}{\theta^2},$$

d'où on déduit de la Proposition 4.2 que

$$\dot{C}_\theta(u) = \theta^{-1} C_\theta(u)\left\{\frac{\ell_\theta(u)}{m_\theta(u)} - \log C_\theta(u)\right\},$$

où

$$\ell_\theta(u) = \sum_{j=1}^{d} u_j^{-\theta}\log u_j \quad et \quad m_\theta(u) = \sum_{j=1}^{d} u_j^{-\theta}.$$

Il s'ensuit que l'hypothèse 3 est satisfaite pour tout $\theta \in (0,\infty)$.

Sous les hypothèses 1–3, on déduit de la Proposition 4.1 que la statistique d'adéquation

$$V_n = \int_{(0,1)^d} \{\mathbb{C}_n(u)\}^2 \, \mathrm{d}C_{\hat{\theta}_n}(u)$$

converge en loi vers

$$V = \int_{(0,1)^d} \{\mathbb{C}(u)\}^2 \, \mathrm{d}C_\theta(u).$$

En développant le carré dans l'expression pour V_n et en utilisant la forme de la copule empirique décrite en 4.1, on montre que

$$\begin{aligned} V_n &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} S(C_{\theta_n})\left(\frac{R_{i1}\vee R_{j1}}{n}, \ldots, \frac{R_{id}\vee R_{jd}}{n}\right) \\ &\quad -2\sum_{i=1}^{n} I_{d,\hat{\theta}_n}\left(\frac{R_{i1}}{n}, \ldots, \frac{R_{id}}{n}\right) + nI^2_{d,\hat{\theta}_n}, \end{aligned}$$

où la fonctionnelle S associe à une fonction de dépendance C sa copule de survie définie par $S(C)(u) = \mathrm{P}(U > u)$, avec $U \sim C$, alors que

$$I_{d,\theta}(a) = \int_{a_1}^{1} \cdots \int_{a_d}^{1} C_\theta(u)\,\mathrm{d}C_\theta(u) \quad \text{et} \quad I^2_{d,\theta} = \int_0^1 \cdots \int_0^1 C^2_\theta(u)\,\mathrm{d}C_\theta(u).$$

Des expressions explicites pour $I_{d,\theta}$ et $I^2_{d,\theta}$ s'obtiennent pour la famille de Clayton. Elles sont présentées dans le tableau 4.2 qui suit.

Table 4.2: Valeurs de $I_{d,\theta}$ et $I^2_{d,\theta}$ pour la famille de Clayton

	$I_{d,\theta}(a)$	$I^2_{d,\theta}$
$d = 2$	$\frac{\theta+1}{2(\theta+2)} S(C^2_\theta)(a)$	$\frac{\theta+1}{3(\theta+3)}$
$d = 3$	$\frac{2\theta+1}{4(\theta+2)} S(C^2_\theta)(a)$	$\frac{(\theta+1)(2\theta+1)}{3(\theta+3)(2\theta+3)}$
$d = 4$	$\frac{4\theta+1}{8(\theta+2)} S(C^2_\theta)(a)$	$\frac{(2\theta+1)(3\theta+1)}{9(\theta+3)(2\theta+3)}$
$d = 5$	$\frac{8\theta+1}{16(\theta+2)} S(C^2_\theta)(a)$	$\frac{(2\theta+1)(3\theta+1)(4\theta+1)}{9(\theta+3)(2\theta+3)(4\theta+3)}$

Ces calculs permettent d'étudier la puissance de V_n sous l'hypothèse nulle que la copule d'une population appartient à la famille de Clayton, en utilisant le *bootstrap* paramétrique décrit au chapitre 2. Dans l'étude de puissance présentée ici, l'algorithme décrit à l'annexe B est utilisé pour simuler des observations de la copule de Clayton. Le paramètre θ est estimé par $\tau^{-1}(\tau_n)$, où $\tau(\theta)$ est le tau de Kendall multivarié défini à l'équation (2.7) et τ_n est sa version empirique. Le tableau 4.3 fournit l'information à ce sujet.

Table 4.3: Tau de Kendall et inversion du tau de Kendall pour la copule multivariée de Clayton de dimensions $d = 2, 3, 4, 5$

d	$\tau(\theta)$	$\tau^{-1}(x)$
2	$\frac{\theta}{\theta+2}$	$\frac{2x}{1-x}$
3	$\frac{\theta}{\theta+2}$	$\frac{2x}{1-x}$
4	$\frac{3\theta(7\theta+4)}{7(\theta+2)(3\theta+2)}$	$\frac{28x-6+2\sqrt{49x^2+63x+9}}{21(1-x)}$
5	$\frac{\theta(9\theta+4)}{3(\theta+2)(3\theta+2)}$	$\frac{12x-2+2\sqrt{9x^2+15x+1}}{9(1-x)}$

Le tableau 4.4 contient les résultats d'une simulation visant à estimer la puissance des statistiques S_n, T_n et V_n sous trois contre-hypothèses multivariées, pour les dimensions $d = 2$ et $d = 3$.

Table 4.4: Puissance, basée sur 10 000 répétitions, des statistiques d'adéquation S_n, T_n et V_n pour vérifier l'hypothèse nulle d'une loi sous-jacente Clayton de dimensions $d = 2$ et $d = 3$ lorsque $n = 250$

Alternative		Dimension du vecteur					
Modèle	Valeur	$d = 2$			$d = 3$		
	de τ	S_n	T_n	V_n	S_n	T_n	V_n
Clayton	0.2	4.6	5.5	5.7	4.0	4.2	5.6
	0.4	4.2	4.7	4.7	4.8	4.7	5.6
	0.6	4.6	4.2	4.3	4.3	4.9	5.2
	0.8	4.7	4.4	2.9	3.9	4.4	4.3
Frank	0.2	47.1	38.9	8.4	3.0	2.7	14.4
	0.4	97.1	88.4	37.4	47.6	42.0	60.0
	0.6	99.9	98.2	92.2	98.9	97.0	99.4
	0.8	100.0	96.2	100.0	100.0	100.0	100.0
Normale	0.2	33.3	23.7	54.8	38.7	30.5	81.3
équicorrélée	0.4	86.9	71.6	97.4	93.5	81.9	100.0
	0.6	99.1	92.1	99.9	99.7	95.4	100.0
	0.8	100.0	98.0	100.0	100.0	98.8	100.0

On remarque que V_n domine les deux autres statistiques pour les contre-hypothèses de Frank et normales trivariées. De plus, V_n présente un avantage marqué sur ses concurrentes lorsque le modèle de dépendance sous-jacent est la copule gaussienne bivariée et trivariée.

4.4 Test de normalité multivariée

Soient N, la fonction de répartition d'une variable $\mathcal{N}(0,1)$ et H_R, la fonction de répartition multivariée d'un vecteur aléatoire normal à d dimensions de matrice de corrélation R définie positive et telle que $R_{jj} = 1$ pour tout $j \in \{1, \dots, d\}$. La forme de l'unique copule associée au modèle multivarié normal est décrite à l'équation (1.3).

La structure de dépendance de ce modèle est complètement déterminée par les valeurs de la matrice R, à savoir

$$R_{jk} = \mathrm{cor}\left\{N^{-1}(U_j),\, N^{-1}(U_k)\right\}, \quad 1 \leq j, k \leq d,$$

où la loi jointe de $U_1, \dots, U_p$ est C_R. Pour estimer les paramètres R_{jk}, $j \neq k$, on pourrait utiliser la statistique de van der Waerden définie par

$$r_{VdW} = \sum_{i=1}^{n} \Phi^{-1}\left(\frac{A_i}{n+1}\right) \Phi^{-1}\left(\frac{B_i}{n+1}\right) \Bigg/ \sum_{i=1}^{n} \left\{\Phi^{-1}\left(\frac{i}{n+1}\right)\right\}^2,$$

où $\{A_i, 1 \leq i \leq n\}$ et $\{B_i, 1 \leq i \leq n\}$ sont deux ensembles de rangs. Donc, pour estimer R_{jk} pour une certaine paire (j,k), il s'agit de prendre $A_i = nF_{nj}(X_{ij})$ et $B_i = nF_{nk}(X_{ik})$. Selon un résultat de Klaassen & Wellner (1997), ce choix s'avère optimal pour de grands échantillons lorsque les observations proviennent d'une loi dont la copule sous-jacente est normale, puisque cet estimateur atteint la borne inférieure de Cramér–Rao, à savoir $(1 - R_{jk}^2)^2$, et est asymptotiquement sans biais.

L'estimateur de van der Waerden semble également posséder de belles propriétés pour des échantillons à tailles finies, tel qu'illustré au tableau suivant.

En effet, on voit que l'erreur quadratique moyenne de r_{VdW} est inférieure à celle de deux estimateurs de rangs naturels, basés respectivement sur le rho de Spearman r_S et le tau de Kendall τ_n. Précisément, les deux estimateurs concurrents qui ont été considérés sont

$$r^{(1)} = 2\sin\left(\frac{\pi}{6}\, r_S\right) \quad \text{et} \quad r^{(2)} = \sin\left(\frac{\pi}{2}\, \tau_n\right).$$

Ces transformations appliquées à r_S et τ_n reflètent le lien entre les versions théoriques de ces mesures d'association et le coefficient de corrélation du modèle normal bivarié.

Table 4.5: Estimation, basée sur 10 000 répétitions, de n fois l'erreur quadratique moyenne pour trois estimateurs non paramétriques du coefficient de corrélation pour la copule normale bivariée de coefficient de corrélation ρ

	$n = 100$			$n = 250$		
ρ	r_{VdW}	$r^{(1)}$	$r^{(2)}$	r_{VdW}	$r^{(1)}$	$r^{(2)}$
0.0	1.0124	1.1054	1.1261	1.0036	1.1017	1.1058
0.2	0.9514	1.0438	1.0554	0.9180	1.0152	1.0180
0.4	0.7302	0.8183	0.8148	0.7480	0.8353	0.8283
0.6	0.4376	0.4995	0.4857	0.4307	0.5017	0.4838
0.8	0.1560	0.1784	0.1639	0.1444	0.1749	0.1608

La proposition suivante, dont la démonstration est reportée à l'Annexe B, donne la distribution asymptotique exacte de $W_n = \sqrt{n}(R_n - R)$, où R_n est une matrice $d \times d$ empirique dont l'élément en position (j, k) est l'estimateur de R_{jk} basé sur la statistique de van der Waerden. Ce résultat constitue une généralisation à d dimensions d'un résultat de Klaassen & Wellner (1997).

Proposition 4.3. *Si R_n est une matrice empirique de van der Waerden basée sur un échantillon $X_1, \ldots, X_n$ de vecteurs i.i.d. dont la copule est C_R, alors W_n converge en loi vers une matrice gaussienne centrée W, où*

$$\begin{aligned} \operatorname{cov}(W_{jk}, W_{\ell q}) &= R_{j\ell}R_{kq} + R_{jq}R_{k\ell} + \frac{1}{2}\, R_{jk}R_{\ell q}\left(R_{j\ell}^2 + R_{jq}^2 + R_{k\ell}^2 + R_{kq}^2\right) \\ &\quad - R_{jk}R_{j\ell}R_{jq} - R_{jk}R_{k\ell}R_{kq} - R_{j\ell}R_{k\ell}R_{\ell q} - R_{jq}R_{kq}R_{\ell q}. \end{aligned}$$

En particulier, on a $\operatorname{var}(W_{jk}) = \left(1 - R_{jk}^2\right)^2$.

En regard de ce résultat, on peut donc envisager un test d'adéquation pour la copule normale, dont la structure de corrélation serait estimée par une matrice de van der Waerden empirique. Par exemple, en autant que les hypothèses 1 et 3 soient satisfaites, on s'assure de la convergence de la fonctionnelle de Kolmogorov–Smirnov

$$W_n = \sqrt{n} \sup_{u \in (0,1)^d} \left| C_n(u) - C_{R_n}(u) \right|.$$

Lorsque $d = 2$, on vérifie que

$$\frac{\partial}{\partial u}\, C_\rho(u,v) = N\left\{ \frac{N^{-1}(v) - \rho N^{-1}(u)}{\sqrt{1-\rho^2}} \right\}$$

et que

$$\begin{aligned} &\dot{C}_\rho(u,v) \\ &\quad = \int_{-\infty}^{N^{-1}(u)} \int_{-\infty}^{N^{-1}(v)} h_\rho(s,t) \left\{ \frac{\rho(1-\rho^2) - \rho(s^2+t^2) + (1-\rho^2)st}{(1-\rho^2)^2} \right\} \mathrm{d}t\, \mathrm{d}s, \end{aligned}$$

ce qui permet de satisfaire les hypothèses 1 et 3 pour $\rho \in [0,1)$.

CHAPITRE 5

LOCAL EFFICIENCY OF A CRAMÉR–VON MISES TEST OF INDEPENDENCE

Résumé

Dans les deux chapitres qui suivent, le comportement asymptotique local de tests pour l'indépendance entre des variables aléatoires est étudié. Le cas bivarié est d'abord traité dans le présent chapitre, qui est constitué d'un article accepté pour publication. On y étudie le test de rangs pour l'indépendance proposé par Deheuvels, qui est basé sur une fonctionnelle de Cramér–von Mises du processus de copule empirique. En utilisant un résultat général sur la distribution asymptotique de ce processus sous des alternatives contiguës, la courbe de puissance locale du test de Deheuvels est calculée dans le cas bivarié et comparée à celle de procédures concurrentes basées sur des statistiques linéaires de rangs. La formule d'inversion de Gil-Pelaez est ensuite employée afin d'effectuer des comparaisons additionnelles en terme d'une extension naturelle de la mesure d'efficacité relative asymptotique de Pitman.

Abstract

Deheuvels proposed a rank test of independence based on a Cramér–von Mises functional of the empirical copula process. Using a general result on the asymptotic distribution of this process under sequences of contiguous alternatives, the local power curve of Deheuvels' test is computed in the bivariate case and compared to that of competing procedures based on linear rank statistics. The Gil-Pelaez inversion formula is used to make additional comparisons in terms of a natural extension of Pitman's measure of asymptotic relative efficiency.

5.1 Introduction

Many procedures have been proposed to test whether two random characters X and Y are independent. The classical approach is based on Pearson's correlation coefficient, but its lack of robustness to outliers and departures from normality eventually led researchers to consider alternative nonparametric procedures.

The most commonly used rank tests of independence—those of Savage, Spearman and van der Waerden in particular—rely on linear rank statistics, which may be conveniently written in the form

$$S_n^J = \frac{1}{n} \sum_{i=1}^{n} J\left(\frac{R_i}{n+1}, \frac{S_i}{n+1}\right) - \bar{J}_n, \tag{5.1}$$

where $J : (0,1)^2 \to \mathbb{R}$ is a score function,

$$\bar{J}_n = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} J\left(\frac{i}{n+1}, \frac{j}{n+1}\right)$$

is a centering factor, and $(R_1, S_1), \ldots, (R_n, S_n)$ are the pairs of ranks associated with a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from some population with bivariate cumulative distribution function $H(x, y)$ and continuous margins $F(x)$ and $G(y)$.

In fact, as shown by Behnen (1971, 1972), essentially all statistics of the form (5.1) yield asymptotically optimal rank tests of independence for suitably selected local alternatives. See Genest & Verret (2005) for a recent account of this literature, which includes major contributions by Bhuchongkul (1964), Shirahata (1974, 1975), and Ciesielska & Ledwina (1983), among others.

In practice, however, it is rarely possible to identify with any precision the form of dependence characterized by a family of alternatives. For this reason, omnibus rank tests seem desirable. Because Sklar (1959) showed that H admits a unique representation

$$H(x,y) = C\{F(x), G(y)\}, \quad x, y \in \mathbb{R}$$

in terms of a copula $C : [0,1]^2 \to [0,1]$, and given that independence between the continuous random variables X and Y occurs if and only if $C(u,v) = C_0(u,v) \equiv uv$ everywhere on its domain, a potentially fruitful rank-based approach to testing independence is rooted in the empirical copula

$$C_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{F_n(X_i) \le u, G_n(Y_i) \le v\},$$

where

$$F_n(x) = \frac{1}{n+1}\sum_{i=1}^{n} \mathbf{1}(X_i \le x) \quad \text{and} \quad G_n(y) = \frac{1}{n+1}\sum_{i=1}^{n} \mathbf{1}(Y_i \le y)$$

are the re-scaled empirical versions of F and G, respectively. Observe that procedures based on C_n are rank-based, as $F_n(X_i) = R_i/(n+1)$ and $G_n(Y_i) = S_i/(n+1)$ for $i \in \{1, \ldots, n\}$.

Deheuvels (1979, 1980, 1981a,b,c) was the first to suggest tests of independence based on a continuous functional measuring the distance between C_n and C_0. This led him to study the weak convergence of the empirical copula process

$$\mathbb{C}_n(u,v) = n^{1/2}\{C_n(u,v) - uv\}$$

and its multivariate extension under the null hypothesis of independence. In particular, this made it possible for him to identify the limiting null distribution of the Cramér–von Mises test statistic based on $\mathbb{C}_n$, although he did not actually compare the performance of tests based on this statistic to any competitor.

In a recent extension of the work of Deheuvels, Genest & Rémillard (2004) report simulations which suggest that Cramér–von Mises statistics are generally more powerful than those based on the classical likelihood ratio statistic assuming normality; see Figures 3–5 in their paper. Because it improves convergence and leads to a simpler formula for the test statistic, the version of the Cramér–von Mises functional they consider is actually based on the centered empirical copula process

$$\tilde{\mathbb{C}}_n(u,v) = n^{1/2}\left\{C_n(u,v) - C_n(u,1)C_n(1,v)\right\},$$

where $C_n(u,1) = C_n(1,u)$ is nothing but the distribution function of a uniform random variable on $\{1/(n+1),\ldots,n/(n+1)\}$. The latter may be defined explicitly by

$$C_n(u,1) = C_n(1,u) = \frac{1}{n}\min\left(n, \lfloor (n+1)u \rfloor\right), \quad 0 \le u \le 1$$

where $\lfloor x \rfloor$ stands for the integer part of x.

Following Genest & Rémillard (2004), therefore, a powerful nonparametric test of independence à la Deheuvels may thus be based on the Cramér–von Mises statistic

$$B_n = \int_{(0,1)^2} \left\{\tilde{\mathbb{C}}_n(u,v)\right\}^2 \mathrm{d}v\,\mathrm{d}u = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} D_n(R_i,R_j)D_n(S_i,S_j),$$

where

$$D_n(s,t) = \frac{2n+1}{6n} + \frac{s(s-1)}{2n(n+1)} + \frac{t(t-1)}{2n(n+1)} - \frac{\max(s,t)}{n+1}.$$

In addition to being simple to compute, this statistic can be simulated easily in order to construct tables of critical values for any fixed sample size n through Monte Carlo methods. Asymptotic critical values for the standard levels may also be found in Table 1 of Genest & Rémillard (2004).

The purpose of this paper is to compare the large-sample performance of standard rank tests of independence to the procedure based on B_n. To this end, the common asymptotic behavior of $\mathbb{C}_n$ and $\tilde{\mathbb{C}}_n$ under contiguous sequences (C_{θ_n}) of parametric alternatives is considered in Section 5.2. The result is then used in Sections 5.3 and 5.4 to derive the asymptotic distribution of S_n^J and B_n under such sequences of alternatives. Examples of calculations are given in Section 5.5.

In Section 5.6, the local asymptotic power curve of the test based on B_n is computed and compared to that of the locally most powerful linear rank statistic, identified by Shirahata (1974, 1975); see also Genest & Verret (2005). A natural extension of Pitman's measure of asymptotic relative efficiency is then used in Section 5.7 to make numerical power comparisons under various families of copula models. Finally, some concluding remarks are made in Section 5.8.

5.2 Asymptotic behavior of $\mathbb{C}_n$

Consider a family (C_θ) of absolutely continuous bivariate copulas indexed by a real parameter $\theta \in \Theta$ in such a way that $C_\theta(u, v)$ is monotone in θ and

$C_{\theta_0}(u,v) = uv$ for all $u, v \in (0,1)$. Let $\delta \in \mathbb{R}$ be such that $\theta_n = \theta_0 + \delta n^{-1/2} \in \Theta$ for n sufficiently large, and suppose that

(i) the density $\partial^2 C_\theta(u,v)/\partial u \partial v = c_\theta(u,v)$ admits a square-integrable, right derivative $\dot{c}_\theta$ at $\theta = \theta_0$ for every fixed $u, v \in (0,1)$, and

$$\lim_{n\to\infty} \int_{(0,1)^2} \left[n^{1/2} \left\{ c_{\theta_n}^{1/2}(u,v) - 1 \right\} - \frac{\delta}{2}\, \dot{c}_{\theta_0}(u,v) \right]^2 \mathrm{d}v \, \mathrm{d}u = 0;$$

(ii) for every $u, v \in (0,1)$, the following identity holds:

$$\dot{C}_{\theta_0}(u,v) = \lim_{\theta\to\theta_0} \frac{\partial}{\partial\theta} C_\theta(u,v) = \int_0^u \int_0^v \dot{c}_{\theta_0}(s,t)\, \mathrm{d}t \, \mathrm{d}s.$$

Let also Q_n denote the joint distribution of a random sample $(X_{n1}, Y_{n1}), \ldots, (X_{nn}, Y_{nn})$ from distribution $C_{\theta_n}\{F(x), G(y)\}$, and denote by P_n the joint distribution of the same sample under independence. As can be deduced from Lemma 3.10.11 of van der Vaart (1996), Condition (i) is sufficient to ensure the contiguity of Q_n with respect to P_n. More precisely, if (F_n) is any sequence of sample-based events such that $P_n(F_n) \to 0$ as $n \to \infty$, then $Q_n(F_n) \to 0$, as $n \to \infty$.

Under these assumptions, the asymptotic behavior of the process $\mathbb{C}_n$ may be characterized as follows.

Proposition 5.1. *Under Conditions* (i)–(ii)*, the sequence of empirical rank processes* $\mathbb{C}_n = n^{1/2}(C_n - C_{\theta_0})$ *converges weakly in* $\mathbf{D}([0,1]^2)$*, under* Q_n*, to a continuous Gaussian limit* $\mathbb{C} + \delta\dot{C}_{\theta_0}$*, where* $\mathbb{C}$ *is a continuous centered normal process such that* $\operatorname{cov}\{\mathbb{C}(u,v), \mathbb{C}(u',v')\} = \gamma(u,u')\gamma(v,v')$*, with* $\gamma(s,t) = \min(s,t) - st$*.*

Proof. Write $U_{ni} = F(X_{ni})$, $V_{ni} = G(Y_{ni})$, and introduce

$$\Phi_n(u) = \frac{1}{n+1}\sum_{i=1}^{n} \mathbf{1}\left(U_{ni} \le u\right) \quad \text{and} \quad \Psi_n(v) = \frac{1}{n+1}\sum_{i=1}^{n} \mathbf{1}\left(V_{ni} \le v\right).$$

Let also

$$\mathbb{A}_n(u,v) = n^{-1/2}\sum_{i=1}^{n} \left\{\mathbf{1}\left(U_{ni} \le u, V_{ni} \le v\right) - uv\right\}.$$

Then

$$\mathbb{C}_n(u,v) = \mathbb{A}_n\left\{\Phi_n^{-1}(u), \Psi_n^{-1}(v)\right\} + n^{1/2}\left\{\Phi_n^{-1}(u)\Psi_n^{-1}(v) - uv\right\}. \tag{5.2}$$

Under Condition (i), it follows from Theorem 3.10.12 of van der Vaart (1996) that, under Q_n, the sequence $(\mathbb{A}_n)$ of processes converges in $\mathbf{D}\left([0,1]^2\right)$ to a continuous Gaussian limit of the form $\mathbb{A} + \delta\dot{C}_{\theta_0}$, where $\dot{C}_{\theta_0}$ is defined as in Condition (ii).

In particular, under Q_n, $\mathbb{A}_n(u,1) = n^{1/2}\left\{\Phi_n(u) - u\right\}$ converges in $\mathbf{D}\left([0,1]\right)$ to $\mathbb{A}(u,1) + \delta\dot{C}_{\theta_0}(u,1)$, and the latter reduces to $\mathbb{A}(u,1)$, since

$$\dot{C}_{\theta_0}(u,1) = \lim_{\theta\to\theta_0} \frac{C_\theta(u,1) - C_{\theta_0}(u,1)}{\theta - \theta_0} = \lim_{\theta\to\theta_0} \frac{u-u}{\theta-\theta_0} = 0.$$

Thus, using identities (11) and (12) in Chapter 3 of Shorack & Wellner (1984), one may deduce that

$$\sup_{u\in[0,1]} \left|\Phi_n(u) - u\right| = \sup_{u\in[0,1]} \left|\Phi_n^{-1}(u) - u\right|$$

tends to zero in probability, whence it follows that $n^{1/2}\left\{\Phi_n^{-1}(u) - u\right\}$ converges in $\mathbf{D}\left([0,1]\right)$ to $-\mathbb{A}(u,1)$. Likewise, $\sup_{v\in[0,1]}\left|\Psi_n^{-1}(v) - v\right|$ tends to zero in probability, and $n^{1/2}\left\{\Psi_n^{-1}(v) - v\right\}$ converges in $\mathbf{D}\left([0,1]\right)$ to $-\mathbb{A}(1,v)$.

Writing the second summand in (5.2) in the alternative form

$$n^{1/2}\left\{\Phi_n^{-1}(u) - u\right\}\Psi_n^{-1}(v) + un^{1/2}\left\{\Psi_n^{-1}(v) - v\right\},$$

one may thus conclude that under Q_n, the empirical process $\mathbb{C}_n$ converges in $\mathbf{D}\left([0,1]^2\right)$ to $\mathbb{C} + \delta\dot{C}_{\theta_0}$, where

$$\mathbb{C}(u,v) = \mathbb{A}(u,v) - v\mathbb{A}(u,1) - u\mathbb{A}(1,v),$$

whose covariance structure is as given in the statement of the proposition. ◇

5.3 Asymptotic behavior of S_n^J

Henceforth, $J : (0,1)^2 \to \mathbb{R}$ is called a score function if it is right-continuous, square-integrable and quasi-monotone, i.e., $J(u',v') - J(u',v) - J(u,v') + J(u,v) \geq 0$ for all $u \leq u'$ and $v \leq v'$. Under these standard conditions, which are met in all classical cases, Quesada-Molina (1992) showed that if (U_i, V_i) is distributed as copula C_i, then

$$\mathrm{E}\left\{J(U_1,V_1) - J(U_2,V_2)\right\} = \int_{(0,1)^2} \left\{C_1(s,t) - C_2(s,t)\right\} \mathrm{d}J(s,t),$$

provided $\mathrm{E}\left\{|J(U_i,V_i)|\right\} < \infty$ for $i = 1,2$. Using this result, one may then reexpress the linear rank statistic S_n^J, defined by (5.1), as

$$n^{1/2}S_n^J = \int_{(0,1)^2} \tilde{\mathbb{C}}_n(u,v)\,\mathrm{d}J(u,v).$$

Since

$$\sup_{u\in[0,1]} |C_n(u,1) - u| \leq \frac{1}{n},$$

$\mathbb{C}_n$ and $\tilde{\mathbb{C}}_n$ obviously have the same limiting behavior under the conditions of Proposition 5.1. Thus for any closed interval $K \subset (0,1)^2$, one has

$$\int_K \tilde{\mathbb{C}}_n(u,v)\,\mathrm{d}J(u,v) \rightsquigarrow \int_K \mathbb{C}(u,v)\,\mathrm{d}J(u,v) + \delta \int_K \dot{C}_{\theta_0}(u,v)\,\mathrm{d}J(u,v),$$

where $\rightsquigarrow$ denotes convergence in law. A technical argument described in Appendix C then implies that $n^{1/2}S_n^J$ converges in law to

$$\mathbb{S}^J = \int_{(0,1)^2} \mathbb{C}(u,v)\,\mathrm{d}J(u,v) + \delta \int_{(0,1)^2} \dot{C}_{\theta_0}(u,v)\,\mathrm{d}J(u,v),$$

under the additional condition

(iii) $\displaystyle\int_{(0,1)^2} \left|\dot{C}_{\theta_0}(u,v)\right| \mathrm{d}J(u,v) < \infty.$

This finding may be summarized as follows.

Proposition 5.2. *Under Conditions* (i)–(iii), $n^{1/2}S_n^J$ *is asymptotically normal, under* Q_n*. Its mean and variance are respectively given by* $\mathrm{E}(\mathbb{S}^J) = \delta\mu_J$ *and* $\mathrm{var}(\mathbb{S}^J) = \sigma_J^2$*, where*

$$\mu_J = \int_{(0,1)^2} \dot{C}_{\theta_0}(u,v)\,\mathrm{d}J(u,v)$$

and

$$\sigma_J^2 = \int_{(0,1)^4} \gamma(u,u')\gamma(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v') = \int_{(0,1)^2} \left\{\tilde{J}(u,v)\right\}^2 \mathrm{d}v\,\mathrm{d}u,$$

with

$$\tilde{J}(u,v) = J(u,v) - \int_{(0,1)} J(u,t)\,\mathrm{d}t - \int_{(0,1)} J(s,v)\,\mathrm{d}s + \int_{(0,1)^2} J(s,t)\,\mathrm{d}s\,\mathrm{d}t.$$

This is consistent with the results already reported by Genest & Verret (2005) under a different set of conditions.

Remark 5.2. *As can be seen from Table 1 below, many classical linear rank statistics have score functions of the form* $J(u,v) = K_1^{-1}(u)K_2^{-1}(v)$*, where, for* $i = 1, 2$*,* F_i *is a cumulative distribution function with zero mean and finite variance* σ_i^2*. In that case, it follows from Proposition* (5.2) *that*

$$n^{1/2} S_n^J = \int_{\mathbb{R}^2} \tilde{C}_n \left\{K_1(x), K_2(y)\right\} \mathrm{d}y \, \mathrm{d}x,$$

whence $J = \tilde{J}$ *and* $\sigma_J^2 = \sigma_1^2 \sigma_2^2$*, as already reported in Proposition 3.1 of Genest & Rémillard (2004).*

5.4 Asymptotic behavior of B_n

Under the conditions of Proposition 5.1, $\dot{C}_{\theta_0}$ is continuous and bounded on $[0,1]^2$, so that the limiting distribution of B_n under the contiguous sequence (Q_n) is given by

$$\mathbb{B} = \int_{(0,1)^2} \left\{\mathbb{C}(u,v) + \delta \dot{C}_{\theta_0}(u,v)\right\}^2 \mathrm{d}v \, \mathrm{d}u.$$

Now it is well known (see, e.g., Shorack and Wellner, 1984, p. 213) that $\mathbb{C}$ admits a Karhunen–Loève expansion

$$\mathbb{C}(u,v) = \sum_{k,\ell \in \mathbb{N}} \lambda_{k\ell}^{1/2} f_{k\ell}(u,v) Z_{k\ell},$$

where the $Z_{k\ell}$ are mutually independent $\mathcal{N}(0,1)$ random variables, and for all integers $k, \ell \in \mathbb{N} = \{1, 2, \ldots\}$,

$$\lambda_{k\ell} = \frac{1}{k^2 \ell^2 \pi^4} \quad \text{and} \quad f_{k\ell}(u,v) = 2 \sin(k\pi u) \sin(\ell \pi v), \quad u, v \in (0,1).$$

Accordingly, one has

$$\mathbb{B}_0 = \int_{(0,1)^2} \{\mathbb{C}(u,v)\}^2 \,\mathrm{d}v\,\mathrm{d}u = \sum_{k,\ell\in\mathbb{N}} \lambda_{k\ell} Z_{k\ell}^2$$

and hence

$$\mathbb{B} = \sum_{k,\ell\in\mathbb{N}} \lambda_{k\ell} Z_{k\ell}^2 + 2\delta \sum_{k,\ell\in\mathbb{N}} \lambda_{k\ell} I_{k\ell} Z_{k\ell} + \delta^2 I,$$

where

$$I = \int_{(0,1)^2} \left\{\dot{C}_{\theta_0}(u,v)\right\}^2 \,\mathrm{d}v\,\mathrm{d}u$$

and

$$I_{k\ell} = \lambda_{k\ell}^{-1/2} \int_{(0,1)^2} f_{k\ell}(u,v)\dot{C}_{\theta_0}(u,v)\,\mathrm{d}v\,\mathrm{d}u.$$

Letting $\chi_1^2(\nu)$ denote a chi-square random variable with one degree of freedom and non-centrality parameter ν, one may then state the following result.

Proposition 5.3. *Under Conditions* (i)–(ii)*, the limiting distribution of B_n, under Q_n, is given by the weighted sum*

$$\mathbb{B} = \sum_{k,\ell\in\mathbb{N}} \lambda_{k\ell} \left(Z_{k\ell} + \delta I_{k\ell}\right)^2 = \sum_{k,\ell\in\mathbb{N}} \lambda_{k\ell} \chi_1^2\left(\delta^2 I_{k\ell}^2\right)$$

of non-central χ_1^2 random variables which depends on the underlying contiguous family (C_{θ_n}) of copula alternatives only through $\dot{C}_{\theta_0}$ via the formula

$$I_{k\ell} = 2k\ell\pi^2 \int_{(0,1)^2} \sin(k\pi u)\sin(\ell\pi v)\dot{C}_{\theta_0}(u,v)\,\mathrm{d}v\,\mathrm{d}u.$$

Proof. From direct substitution into the integral representation of $\mathbb{B}$ of Parseval's identity, $I = \sum_{k,\ell\in\mathbb{N}} \lambda_{k\ell} I_{k\ell}^2$. ♢

5.5 Examples

Several commonly used families of bivariate copulas satisfy Conditions (i)–(ii). Interestingly, many of them yield the same value for $\dot{C}_{\theta_0}$, up to a constant. The copula models listed, e.g., in the books of Joe (1997), Nelsen (1999) or Drouet-Mari (2001) may thus be clustered into classes whose members all lead to essentially the same asymptotic distribution for B_n. Here are three examples.

Class 1. A simple calculation shows that $\dot{C}_{\theta_0}(u,v) \propto uv(1-u)(1-v)$ for the Ali–Mikhail–Haq, Dabrowska (Oakes and Wang 2003), Farlie–Gumbel–Morgenstern, Frank, and Plackett families of copulas. Note that Condition (iii) holds for any score function J, and that μ_J is proportional to

$$\int_{(0,1)^2} (1-2u)(1-2v)\tilde{J}(u,v)\,\mathrm{d}v\,\mathrm{d}u = \int_{(0,1)^2} (1-2u)(1-2v)J(u,v)\,\mathrm{d}v\,\mathrm{d}u,$$

while

$$I_{k\ell} \propto \begin{cases} \dfrac{32}{k^2\ell^2\pi^4} & \text{if } k \text{ and } \ell \text{ are odd,} \\ 0 & \text{otherwise.} \end{cases}$$

Class 2. For the Clayton and Gumbel–Barnett families, as well as for Model 4.2.10 of Nelsen (1999), one has $\dot{C}_{\theta_0}(u,v) \propto \pm uv \log(u)\log(v)$, and hence

$$I_{k\ell} \propto \pm \frac{2}{k\ell\pi^2}\, SI(k\pi)SI(\ell\pi), \quad \text{where } SI(x) = \int_0^x t^{-1}\sin(t)\,\mathrm{d}t.$$

Class 3. If C_θ is the Gaussian copula and N denotes the cumulative distribution function of a $\mathcal{N}(0,1)$ random variable, then

$$\dot{C}_{\theta_0}(u,v) = N'\left\{N^{-1}(u)\right\} N'\left\{N^{-1}(v)\right\}$$

with $N' = dN(t)/\,\mathrm{d}t$, so that $I_{k\ell} = 2k\ell\pi^2\, g(k)g(\ell)$, where

$$g(m) = \int_{\mathbb{R}} \{N'(t)\}^2 \sin\{m\pi\Phi(t)\}\,\mathrm{d}t.$$

Because of their connection with frailty models, bivariate Archimedean copula models (Nelsen 1999, Chap. 4) are particularly common in practice. They can be expressed in the form

$$C_\theta(u,v) = \psi_\theta^{-1}\left\{\psi_\theta(u) + \psi_\theta(v)\right\}$$

in terms of a generator $\psi_\theta : (0,1] \to [0,\infty)$ which is convex, decreasing, and such that $\psi_\theta(1) = 0$. A simple formula for $\dot{C}_{\theta_0}$ is given next for such models, under the assumption that $\dot{\psi}_\theta(t) = \partial\psi_\theta(t)/\partial\theta$ exists and is continuous in a neighborhood of θ_0. The result extends readily to the multivariate case.

Proposition 5.4. *If (C_θ) is a parametric family of Archimedean copulas whose generators ψ_θ are normalized in such a way that $\psi_\theta(t) \to -\log(t)$ and $\psi'_\theta(t) \to -1/t$ as $\theta \to \theta_0$, then*

$$\dot{C}_{\theta_0}(u,v) = uv\left\{\dot{\psi}_{\theta_0}(uv) - \dot{\psi}_{\theta_0}(u) - \dot{\psi}_{\theta_0}(v)\right\}.$$

Proof. The conclusion obtains by letting $\theta \to \theta_0$ in the expression

$$\begin{aligned}\dot{C}_\theta(u,v) &= \frac{\partial}{\partial\theta}\,\psi_\theta^{-1}(t)\bigg|_{t=\psi_\theta(u)+\psi_\theta(v)} + \left\{\dot{\psi}_\theta(u) + \dot{\psi}_\theta(v)\right\}\frac{\partial}{\partial t}\,\psi_\theta^{-1}(t)\bigg|_{t=\psi_\theta(u)+\psi_\theta(v)} \\ &= -\frac{\dot{\psi}_\theta\{C_\theta(u,v)\}}{\psi'_\theta\{C_\theta(u,v)\}} + \frac{\dot{\psi}_\theta(u) + \dot{\psi}_\theta(v)}{\psi'_\theta\{C_\theta(u,v)\}}\,,\end{aligned}$$

which results from straightforward applications of the Chain Rule and the Inverse Function Theorem. ◊

5.6 Comparisons between tests based on B_n and S_n^J

In addition to characterizing the asymptotic behavior of tests of independence based on B_n or S_n^J, Propositions (5.2) and (5.3) help to delineate the circumstances under which these various procedures might perform best.

5.6.1 Consistency

An advantage of basing a test of independence on B_n is that it is always consistent. Such is not necessarily the case for procedures involving S_n^J. Assume, for instance, that the data arise from the family (C_r) of Student copulas indexed by their "correlation coefficient" r, as is often assumed in financial applications (see Cherubini, Luciano and Vecchiato 2004 and references therein). Note that in this case, C_0 is *not* the independence copula.

Now suppose that J is a score function such that

$$J(u,v) + J(u,1-v) + J(1-u,v) + J(1-u,1-v) = 0$$

for all $u, v \in (0,1)$. Under the latter condition, which is met for several of the classical score functions listed in Table 5.1, one finds $\bar{J}_n = 0$ and

$$\int_{(0,1)^2} J(u,v)\,\mathrm{d}C_0(u,v) = 0 \tag{5.3}$$

whenever this integral exists.

The main result in Chapter 5 of Gänssler & Stute (1987), coupled with Quesada-Molina's identity, implies that

$$n^{1/2} S_n^J \to \tilde{\mathbb{S}}^J = \int_{(0,1)^2} \tilde{\mathbb{C}}(u,v)\,\mathrm{d}J(u,v),$$

where

$$\tilde{C}(u,v) = \tilde{\mathbb{A}}(u,v) - u\tilde{\mathbb{A}}(1,v) - v\tilde{\mathbb{A}}(u,1)$$

and $\mathbb{A}$ is the limiting distribution of the process $n^{1/2}\{C_n(u,v) - C_0(u,v)\}$. In view of (5.3), $\mathbb{S}^J$ is Gaussian with zero mean, so that the test based on this particular S_n^J would be inconsistent, while B_n/n would still converge in probability to

$$\int_{(0,1)^2} \{C_0(u,v) - uv\}^2\,\mathrm{d}v\,\mathrm{d}u > 0.$$

Note, incidentally, that the same inconsistent behavior of S_n^J would hold true for any non-Gaussian, meta-elliptical copula with $r = 0$. See Abdous et al. (2005) for related properties of this large class of copulas.

5.6.2 Asymptotic local power

Additional comparisons between procedures based on B_n and S_n^J can be made through the notion of asymptotic local power function for tests of size α based on these statistics. Letting $z_{\alpha/2} = N^{-1}(1-\alpha/2)$ represent the quantile of order $1-\alpha/2$ of a standard normal random variable Z, and assuming the conditions of Proposition (5.2), one can see that the asymptotic local power of the test based on S_n^J along the sequence (Q_n) of contiguous alternatives is given by

$$\beta_{S^J}(\delta,\alpha) = \lim_{n\to\infty} Q_n\left(\left|n^{1/2}S_n^J\right| > \sigma_J z_{\alpha/2}\right) = \mathrm{P}\left(\left|Z + \delta\mu_J/\sigma_J\right| > z_{\alpha/2}\right).$$

Note that since the mapping $a \mapsto \mathrm{P}(-z_{\alpha/2}-a \leq Z \leq z_{\alpha/2}-a)$ is decreasing in a on $[0,\infty)$, a rank test of size α based on score function J will be preferable to another rank test of the same size based on score function K whenever $|\mu_J/\sigma_J| > |\mu_K/\sigma_K|$. Moreover,

$$\mathrm{ARE}\left(S^J, S^K\right) = \left(\frac{\mu_J/\sigma_J}{\mu_K/\sigma_K}\right)^2,$$

known as Pitman's asymptotic relative efficiency, may be interpreted as the ratio of sample sizes required for the two test statistics to maintain the same level and power along the contiguous sequence (C_{θ_n}) of copula alternatives. Obviously, the index $\mathrm{ARE}(S^J, S^K)$ is the same for any two families (C_θ) and (D_λ) in the same class, i.e., whenever $\dot{C}_{\theta_0} \propto \dot{D}_{\lambda_0}$.

Listed in Table 5.1 are the score functions J of some linear rank statistics S_n^J that satisfy the conditions of Proposition 5.2. Except for two, they are all products of quantile functions, and hence the Remark at the end of Section 5.3 applies to them. The exceptions are the symmetrized versions of the Wilcoxon and Blest (2000) statistics, obtained by taking $J^*(u,v) = J(u,v) + J(v,u)$. (See Genest and Plante 2003 for additional details.)

Table 5.2 gives the value of $\mathrm{ARE}(S^J, S^{J_{\mathrm{opt}}})$ for the various choices of J listed in Table 5.1 and $J_{\mathrm{opt}} \propto \dot{c}_{\theta_0}$ for the three families of copulas considered in Section 5.5. As shown by Genest & Verret (2005), this choice of J_{opt} corresponds to the locally most powerful rank test statistic for the family of alternatives under consideration. The calculation of the ARE for the symmetrized statistics is facilitated by the fact that when $J(u,v) = K_1^{-1}(u)K_2^{-1}(v)$ is a product of quantile functions with mean zero and finite variance, Proposition 5.2

implies that

$$\text{ARE}\left(S^{J^*}, S^J\right) = \frac{2}{1+\rho^2} \geq 1,$$

where $\rho = \text{corr}\{K_1^{-1}(U), K_2^{-1}(U)\}$.

Table 5.1: Score function of some linear rank statistics S_n^J whose expectation vanishes under the null hypothesis of independence

Test statistic	$J(u,v)$
Blest	$\{1-3(1-u)^2\}\,(2v-1)$
Symmetrized Blest	$(3-u-v)\,\{3\,(2u-1)\,(2v-1)-1\}+2$
Exponential	$\{1+\log(1-u)\}\{1+\log(1-v)\}$
Laplace	$\xi(u)\xi(v)$
Savage	$(1+\log u)(1+\log v)$
Spearman	$(2u-1)\,(2v-1)$
van der Waerden	$\Phi^{-1}(u)\Phi^{-1}(v)$
Wilcoxon	$(2u-1)\log\left(\dfrac{v}{1-v}\right)$
Symmetrized Wilcoxon	$(2u-1)\log\left(\dfrac{v}{1-v}\right)+(2v-1)\log\left(\dfrac{u}{1-u}\right)$

with $\xi(u) = 0.5\ \text{sign}(1/2-u)\log\{2\min(u,1-u)\}$

As clearly illustrated in Table 5.2, the performance of a linear rank statistic can vary substantially when it is compared to the locally most powerful rank test of independence within a given class. It is as low as 41.59% for the exponential rank statistic when alternatives belong to the Clayton family, for example, but it reaches 99.07% for the symmetrized version of Wilcoxon's rank statistic in the normal copula model.

Table 5.2: Pitman asymptotic relative efficiency $\text{ARE}(S^J, S^{J_{\text{opt}}})$ of test statistic S_n^J versus the locally most powerful rank test $S_n^{J_{\text{opt}}}$ of independence for local alternatives in the three classes of copulas considered in Section 5.5

	Copula families from		
Test statistic	Class 1	Class 2	Class 3
Blest	$\frac{15}{16} = 0.9375$	$\frac{125}{192} \approx 0.6510$	$\frac{135}{16\pi^2} \approx 0.8549$
Symmetrized Blest	$\frac{30}{31} \approx 0.9677$	$\frac{125}{186} \approx 0.6720$	$\frac{270}{31\pi^2} \approx 0.8825$
Exponential	$\frac{9}{16} = 0.5625$	$\frac{(\pi^2-6)^2}{36} \approx 0.4159$	0.6655
Laplace	$\frac{729}{1024} \approx 0.7119$	0.6615	0.9274
Savage	$\frac{9}{16} = 0.5625$	1.0000	0.6655
Spearman	1.000	$\frac{9}{16} = 0.5625$	$\frac{9}{\pi^2} \approx 0.9119$
van der Waerden	$\frac{9}{\pi^2} \approx 0.9119$	0.6655	1.0000
Wilcoxon	$\frac{9}{\pi^2} \approx 0.9119$	$\frac{\pi^2}{16} \approx 0.6169$	0.9471
Symmetrized Wilcoxon	$\frac{18}{9+\pi^2} \approx 0.9539$	$\frac{\pi^4}{8(\pi^2+9)} \approx 0.6453$	0.9907

In a sense, however, the AREs reported in Table 5.2 are deceptively low. For, it should be borne in mind that while no linear rank test can ever be more efficient that the locally most powerful procedure, identification of the latter is contingent on the exact knowledge of the direction in which departures from independence occur.

To make comparisons with the Cramér–von Mises statistic B_n, one must resort to the following formula of Gil-Pelaez (1951), which states that if X is a random variable with continuous distribution function F and characteristic function $\hat{f}$, then

$$1 - F(x) = \frac{1}{2} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{Im}\left\{ t^{-1} e^{-ixt} \hat{f}(t) \right\} \mathrm{d}t,$$

where $\mathrm{Im}(z)$ denotes the imaginary part of the complex number z.

To use this identity in the present context, proceed as in Imhof (1961) and write

$$\hat{\eta}(t,\mu) = \frac{1}{(1-2it)^{1/2}}\, e^{\left(\frac{i\mu t}{1-2it}\right)} = \frac{1}{(1+4t^2)^{1/4}}\; e^{-\frac{2t\mu^2}{1+4t^2}}\; e^{it\frac{\mu}{1+4t^2}}\; e^{i\arctan(2t)/2}\,.$$

Then call on Proposition 5.3 to see that

$$\hat{f}(t,\delta) = \mathrm{E}\left(e^{it\mathbb{B}}\right) = \prod_{k,\,\ell\,\in\mathbb{N}} \hat{\eta}\left(\lambda_{k\ell}\, t, \delta^2 I_{k\ell}^2\right) = \xi(t)\, e^{-2\delta^2 t^2 \kappa_1(t)}\, e^{i\kappa_2(t) + i\delta^2\kappa_3(t)},$$

where

$$\begin{aligned}
\xi(t) &= \prod_{k,\,\ell\,\in\mathbb{N}} (1+4t^2\lambda_{k\ell}^2)^{-1/4}, \quad \kappa_1(t) = \sum_{k,\,\ell\,\in\mathbb{N}} \lambda_{k\ell}^2 I_{k\ell}^2/(1+4t^2\lambda_{k\ell}^2),\\
\kappa_2(t) &= \frac{1}{2}\sum_{k,\,\ell\,\in\mathbb{N}} \arctan(2t\lambda_{k\ell}), \quad \kappa_3(t) = t\sum_{k,\,\ell\,\in\mathbb{N}} \lambda_{k\ell} I_{k\ell}^2/(1+4t^2\lambda_{k\ell}^2).
\end{aligned}$$

Note that $\xi(t)$ and $t^2\xi(t)$ are integrable, that κ_1 is bounded, that $\kappa_i(t)/t$ is bounded for $i = 2, 3$, and that $\kappa_2(t)/t \to 1/36$ and $\kappa_3(t)/t \to I^2$, as $t \to 0$.

In the light of the Gil-Pelaez formula, one may deduce that

$$\mathrm{P}(\mathbb{B} > x) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \frac{\sin\{\kappa(x,t)\}}{t\zeta(t)}\, \mathrm{d}t, \tag{5.4}$$

where

$$\kappa(x,t) = -\frac{xt}{2} + \frac{1}{2}\sum_{k,\ell\in\mathbb{N}}\left\{\arctan(\lambda_{k\ell}\, t) + \delta^2\,\frac{\lambda_{k\ell} I_{k\ell}^2 t}{1+\lambda_{k\ell}^2\, t^2}\right\}$$

and

$$\zeta(t) = \exp\left(\frac{\delta^2 t^2}{2}\sum_{k,\ell\in\mathbb{N}}\frac{\lambda_{k\ell}^2 I_{k\ell}^2}{1+\lambda_{k\ell}^2\, t^2}\right)\prod_{k,\ell\in\mathbb{N}}\left(1+\lambda_{k\ell}^2 t^2\right)^{1/4}.$$

Accordingly, numerical approximation routines can be used to compute the local power function $\beta_B(\delta,\alpha) = \mathrm{P}(\mathbb{B} > p_\alpha)$ of B_n. The critical values $p_\alpha =$ 0.0469, 0.0592 and 0.0869 correspond to the traditional levels $\alpha = 0.1$, 0.05, and 0.01, respectively.

Figure 5.1 compares graphically the power of the 5%–level rank tests of independence based on the Cramér–von Mises statistic (broken line) and the locally most powerful procedure (solid line) for the three classes of parametric copula alternatives considered in Section 5.5. Panels 1–3 (from left to right) correspond to Classes 1–3, for which the optimal rank tests are based on the Spearman, Savage and van der Waerden statistics, respectively.

The plotted curves are based on a numerical approximation of (5.4) obtained by integrating on $[0, 100]$ and restricting the sum and integral to integers $k, \ell \le 10$, which guaranteed numerical stability within computer accuracy. As the picture highlights, the power of the test based on B_n is generally close to that of the optimal rank statistic S_n^J with $J = \dot{c}_{\theta_0}$. The statistic B_n does best for Class 1 alternatives in the neighborhood of independence; its performance is least impressive for moderate values of δ in Class 2, i.e., dependence models of the Clayton or Gumbel–Barnett variety.

Figure 5.1: Comparative power of two rank statistics used to test independence for alternatives from three different classes of copulas: broken line, Cramér–von Mises statistic; solid line, locally most powerful procedure

5.7 Asymptotic relative efficiency calculations

For score functions J and K, the asymptotic relative efficiency

$$\mathrm{ARE}\left(S^J, S^K\right) = \left(\frac{\mu_J/\sigma_J}{\mu_K/\sigma_K}\right)^2$$

is a natural measure of local power comparison because the asymptotic behavior of the related statistics S_n^J and S_n^K is Gaussian, under the assumptions of Proposition 5.2. However, a more general definition of asymptotic relative efficiency is needed if comparisons must be extended to statistics such as B_n, whose limiting distribution is not normal. Several options exist; see, e.g., Nyblom & Mäkeläinen (1983) and references therein.

The approach pursued here for comparing tests based on statistics T_n and T_n' involves a ratio of the slopes of the power curves in a neighborhood of $\delta = 0$, viz.

$$e\left(T, T'\right) = \lim_{\delta \to 0} \frac{\beta_T(\delta, \alpha) - \alpha}{\beta_{T'}(\delta, \alpha) - \alpha}.$$

This ratio, which is superior to 1 for all T' whenever T is locally most powerful, provides a natural extension of Pitman's efficiency beyond the case of normal statistics. For, suppose that under Q_n, the limiting power function

$\beta_T(\delta, \alpha)$ of tests of size α based on T_n is given by

$$\beta_T(\delta, \alpha) = 1 - N(z_{\alpha/2} - \delta\mu_T/\sigma_T) + N(-z_{\alpha/2} - \delta\mu_T/\sigma_T),$$

where N is the distribution function of the standard Gaussian and N' is the corresponding density. Then

$$\lim_{\delta\to 0} \delta^{-2} \{\beta_T(\delta, \alpha) - \alpha\} = z_{\alpha/2}N'(z_{\alpha/2}) \, (\mu_T/\sigma_T)^2 ,$$

and hence

$$e\,(T, T') = \lim_{\delta\to 0} \frac{\beta_T(\delta, \alpha) - \alpha}{\beta_{T'}(\delta, \alpha) - \alpha} = \text{ARE}\,(T, T') \, .$$

The following proposition characterizes the local behavior of $\beta_B(\delta, \alpha) - \alpha$ at $\delta = 0$ for the Cramér–von Mises statistic B_n. The proof of this result, which uses the Gil-Pelaez representation, is given in Appendix C.

Proposition 5.5. *Under Conditions* (i)–(ii)*, one has*

$$\lim_{\delta\to 0} \delta^{-2}\{\beta_B(\delta, \alpha) - \alpha\} = \sum_{k,\,\ell\,\in\mathbb{N}} \lambda_{k\ell} I_{k\ell}^2 \, h_{k\ell}(p_\alpha),$$

where $h_{k\ell}$ is a density whose associated characteristic function

$$\frac{\hat{f}(t, 0)}{1 - 2i\lambda_{k\ell}\, t} = (1 - 2i\lambda_{k\ell}\, t)^{-1} \prod_{q,\,r\,\in\mathbb{N}} (1 - 2i\lambda_{qr}\, t)^{-1/2}$$

is that of $\mathbb{B}_0 + \lambda_{k\ell}\, \chi_2^2$, in which the summands are taken to be independent.

Finally, note that

$$h_{k\ell}(x) = \frac{1}{\pi} \int_0^\infty (1 + 4\lambda_{k\ell}^2\, t^2)^{-1/2} \xi(t) \cos\{\kappa_2(t) + \arctan(\lambda_{k\ell}) - tx\} \, \mathrm{d}t,$$

from which it is possible to conclude that

$$\begin{aligned} e\left(B,\, S^J\right) &= \lim_{\delta\to 0}\delta^{-2}\left\{\beta_B(\delta,\alpha)-\alpha\right\}\Big/\lim_{\delta\to 0}\delta^{-2}\left\{\beta_{S^J}(\delta,\alpha)-\alpha\right\} \\ &= \frac{1}{z_{\alpha/2}N'(z_{\alpha/2})\ (\mu_J/\sigma_J)^2}\sum_{k,\,\ell\in\mathbb{N}}\lambda_{k\ell}I_{k\ell}^2\,h_{k\ell}(p_\alpha)\,, \end{aligned}$$

whenever the score function also satisfies Condition (iii).

The local asymptotic relative efficiencies $e\left(B, S^{J_{\rm opt}}\right)$ of B_n with respect to the locally most powerful statistic $S_n^{J_{\rm opt}}$ for copulas from Classes 1–3 are presented in Table 5.3. These numerical approximations were obtained by integrating on $[0, 1500]$ by the trapezoidal rule (with a mesh of $1/2500$) and restricting the sum and product to terms with integers $k, \ell \leq 10$ in the Gil-Pelaez formula.

Table 5.3: Local asymptotic relative efficiency of B_n with respect to the locally most powerful statistic $S^{J_{\rm opt}}$ for three classes of copulas

	Copula families from		
Level	Class 1	Class 2	Class 3
1%	0.8337	0.4229	0.6961
5%	0.8122	0.4181	0.6791
10%	0.8380	0.4386	0.7019

Comparisons involving any other linear rank statistic S_n^J in Table 5.1 may be made easily since

$$e\left(B, S^J\right) = \frac{e\left(B,\, S^{J_{\rm opt}}\right)}{e\left(S^J, S^{J_{\rm opt}}\right)}\,.$$

In conformance with Figure 5.1, B_n is seen to do quite well against the locally most powerful nonparametric test of independence for Class 1 alternatives. Its performance is somewhat worse for Class 3 Gaussian alternatives, and more questionable for Class 2 alternatives, namely the Clayton and Gumbel–Barnett copulas. A rationale for this phenomenon is still lacking.

5.8 Conclusion

Because they allow analysts to model dependence separately from the margins, copulas provide a handy (and increasingly popular) way of constructing alternatives to independence in multivariate contexts. This paper identifies conditions under which a family of copulas gives rise to a contiguous sequence of alternatives. The asymptotic behavior of the empirical copula process is characterized under alternatives of this sort. This leads to a computable expression for the limiting local power of a bivariate Cramér–von Mises statistic originally suggested by Deheuvels, and to meaningful asymptotic relative efficiency comparisons with various linear rank tests of independence.

In addition to being easy to implement, Deheuvels' test based on B_n is always consistent. The numerical comparisons reported in Figure 5.1 and Table 5.3 also show that as an *omnibus* procedure, it generally holds up its power reasonably well against the *model-specific* locally most powerful rank-based test. Considering that the latter test may not be consistent if the alternatives have not been specified correctly, the test based on B_n certainly represents a viable solution, if not an ideal one. Its mitigated success in reproducing the optimal power is obviously a function of the type of departure from independence embodied in the family of local alternatives. Just what aspect of association is at stakes seems hard to pin down, however.

CHAPITRE 6

ASYMPTOTIC LOCAL EFFICIENCY OF CRAMÉR–VON MISES TYPE TESTS FOR MULTIVARIATE INDEPENDENCE

Résumé

Dans cet article, le cas multivarié sera traité, par opposition au cas bivarié. Ceci permettra d'étendre d'une certaine façon les résultats du chapitre précédent. Deheuvels (1981a,b,c) et Genest & Rémillard (2004) ont montré que des tests basés sur les rangs puissants pour l'indépendance multivariée peuvent être bâtis à partir de combinaisons de statistiques de Cramér–von Mises asymptotiquement indépendantes déduites d'une décomposition de Möbius du processus de copule empirique. Un résultat sur le comportement limite de ce processus sous des suites de contre-hypothèses contiguës est utilisé pour donner une représentation de la distribution limite de telles statistiques de test et pour calculer leur efficacité dans un voisinage de l'indépendance. Les courbes de puissance locale de ces statistiques, de même

que l'efficacité relative asymptotique de quelques paires de tests, sont calculées pour des classes de contre-hypothèses fondées sur des familles de copules familières.

Abstract

Deheuvels (1981a,b,c) and Genest & Rémillard (2004) have shown that powerful rank tests of multivariate independence can be based on combinations of asymptotically independent Cramér–von Mises statistics derived from a Möbius decomposition of the empirical copula process. A result on the large-sample behavior of this process under contiguous sequences of alternatives is used here to give a representation for the limiting distribution of such test statistics and to compute their relative local asymptotic efficiency. Local power curves and asymptotic relative efficiencies are compared under familiar classes of copula alternatives.

6.1 Introduction

In a seminal paper concerned with testing the null hypothesis of independence between the $d \geq 2$ components of a multivariate vector with continuous distribution H and marginals $F_1, \ldots, F_d$, Blum et al. (1961) investigated the use of a Cramér–von Mises statistic derived from the process

$$\mathbb{H}_n(x) = \sqrt{n}\left\{H_n(x) - \prod_{j=1}^{d} F_{j,n}(x_j)\right\}, \quad x = (x_1, \ldots, x_d) \in \mathbb{R}^d$$

that measures the difference between the empirical distribution function (EDF) H_n of H and the product of the marginal EDFs $F_{j,n}$ associated with the components of the random vector. As Hoeffding (1948) had already noted, the asymptotic distribution of this test is generally not tractable, and hence tables of critical values are required for its use. Such tables were provided by Blum et al. (1961) themselves in the case $d = 2$, and were later expanded to $d \geq 3$ by Csörgő (1979) and Cotterill & Csörgő (1982, 1985), based on strong approximations of $\mathbb{H}_n$. See also Jing & Zhu (1996) for a bootstrap approach.

Despite this work and the anticipation that the Cramér–von Mises statistic

$$\int \mathbb{H}_n^2 \, \mathrm{d}H_n$$

should be powerful, most subsequent research focussed on the case $d = 2$, where alternative tests (typically based on moment characterizations of independence) were proposed by Feuerverger (1993), Shih & Louis (1996b), Gieser & Randles (1997) and Kallenberg & Ledwina (1999), among others.

Curiously, the literature seems to have largely ignored a suggestion of Blum et al. (1961) to circumvent the inconvenience caused by the complex nature of the limiting distribution of $\mathbb{H}_n$. To be specific, let $X_1 = (X_{11}, \ldots, X_{1d}), \ldots, X_n = (X_{n1}, \ldots, X_{nd})$ be a random sample from distribution H, and for arbitrary $A \subset \mathcal{S}_d = \{1, \ldots, d\}$ with $|A| > 1$, consider the empirical process

$$G_{A,n}(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \prod_{j \in A} \{\mathbf{1}\,(X_{ij} \le x_j) - F_{j,n}\,(x_j)\}.$$

Using Möbius' inversion formula, stated in Section 6.3, Blum, Kiefer and Rosenblatt showed that $\mathbb{H}_n$ may be conveniently expressed as

$$\mathbb{H}_n(x) = \sum_{A \subset \mathcal{S}_d,\, |A|>1} G_{A,n}(x) \prod_{j \in \mathcal{S}_d \setminus A} F_{j,n}(x_j).$$

Although their paper only discussed the case $d = 3$, these authors claimed (and this was later confirmed by Dugué (1975) that under the hypothesis of independence, $G_{A,n}$ converges weakly to a continuous centered Gaussian process with covariance function

$$\text{cov}_A(x, y) = \prod_{j \in A} [\min\{F_j(x_j), F_j(y_j)\} - F_j(x_j)F_j(y_j)]$$

whose eigenvalues, given by

$$\frac{1}{\pi^{2|A|}(i_1 \cdots i_{|A|})^2}\ , \quad i_1, \ldots, i_{|A|} \in \mathbb{N} = \{1, 2, \ldots\}$$

may be deduced from the Karhunen–Loève decomposition of the Brownian bridge. More importantly still, Blum et al. (1961) and Dugué (1975) pointed out that the processes $G_{A,n}$ and $G_{A',n}$ are mutually independent asymptotically whenever $A \ne A'$, so that Cramér–von Mises statistics based on the individual $G_{A,n}$'s could be combined to construct suitable tests against independence.

An obvious limitation of tests based on this approach, however, is the dependence of the asymptotic null distribution of the $G_{A,n}$'s on the marginals of H. To alleviate this problem, Deheuvels (1981a) suggested that the original observations $X_1, \ldots, X_n$ be replaced by their associated rank vectors $R_1 = (R_{11}, \ldots, R_{1d}), \ldots, R_n = (R_{n1}, \ldots, R_{nd})$, where

$$R_{ij} = \sum_{\ell=1}^{n} \mathbf{1}(X_{\ell j} \leq X_{ij}), \quad 1 \leq i \leq n, \quad 1 \leq j \leq d.$$

Deheuvels then went on to characterize the asymptotic null behavior of a Möbius decomposition of the copula process

$$\mathbb{C}_n(u) = \sqrt{n} \left\{ C_n(u) - \prod_{j=1}^{d} u_j \right\}, \tag{6.1}$$

where

$$C_n(u_1, \ldots, u_d) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} \mathbf{1}(R_{ij} \leq n u_j) \tag{6.2}$$

is an estimation of the unique copula C (Sklar 1959) defined implicitly by

$$C\{F_1(x_1), \ldots, F_d(x_d)\} = H(x_1, \ldots, x_d), \quad x_1, \ldots, x_d \in \mathbb{R}.$$

The latter reduces to $C(u) = u_1 \cdots u_d$ under the null hypothesis of independence. Deheuvels (1981a,b,c) thus proposed that this hypothesis be tested using Cramér–von Mises and Kolmogorov–Smirnov statistics based on a decomposition of $\mathbb{C}_n$, but limited himself to the determination of the asymptotic null distribution.

Recently, Genest & Rémillard (2004) showed how to compute the quantiles for the finite-sample and asymptotic null distribution of Deheuvel's proposed statistics based on $\mathbb{C}_n$. Furthermore, they investigated how the $2^d - d - 1$ statistics derived from the rank analogues of the $G_{A,n}$'s could be combined to obtain a global statistic for testing independence, both in the serial and in the non-serial case.

This paper enhances the work of Genest & Rémillard (2004) by comparing the power of Cramér–von Mises tests of independence based on the copula process $\mathbb{C}_n$ and on four different combination recipes for the terms of its Möbius decomposition. To this end, the local asymptotic behavior of the copula process $\mathbb{C}_n$ is characterized in Section 6.2 under a sequence of contiguous alternatives to independence. Examples of contiguous copula alternatives are considered in Section 6.3, where it is shown that the Clayton gamma frailty models share with the equicorrelated Gaussian model the surprising property that the limiting behavior of $\mathbb{G}_{A,n}$ is independent of the sequence of contiguous alternatives for all $A \in \mathcal{S}_d$ with $|A| > 2$. In Section 6.4, some test statistics for the null hypothesis of independence are considered and their asymptotic behavior under contiguous alternatives is specified. Local power functions are then computed in Section 6.5. Finally, Section 6.6 gives additional comparisons between the test statistics using an extension of Pitman's local asymptotic efficiency considered by Genest et al. (2005a) in a bivariate setting.

6.2 Limiting behavior of $\mathbb{C}_n$ under contiguous alternatives

Let $\Theta \subset \mathbb{R}$ be a closed interval and $\mathcal{C} = \{C_\theta,\ \theta \in \Theta\}$ be a given family of copulas that are monotone in θ and for which $\theta_0 \in \Theta$ corresponds to independence. Take $\delta_n \to \delta \in \mathbb{R}$ such that $\theta_n = \theta_0 + \delta_n/\sqrt{n} \in \Theta$ for n large enough.

Let Q_n be the joint distribution function of the random sample

$$\left(X_{11}^{(n)}, \ldots, X_{1d}^{(n)}\right), \ldots, \left(X_{n1}^{(n)}, \ldots, X_{nd}^{(n)}\right)$$

from distribution function $C_{\theta_n}\{F_1(x_1), \ldots, F_d(x_d)\}$ and let C_n be the empirical copula of these observations computed using formula (6.2). Finally, let P_n be the joint distribution of a sample of the same size under the independence distribution $F_1 \times \cdots \times F_d$.

The limiting behavior of the sequence $(\mathbb{C}_n)$ of empirical copula processes will be determined under the following conditions:

(i) C_θ is absolutely continuous and its density $c_\theta(u) = \partial^d C_\theta(u)/\partial u_1 \cdots \partial u_d$ admits a square-integrable, right-derivative $\dot{c}_\theta(u) = \partial c_\theta(u)/\partial\theta$ at $\theta = \theta_0$ for every $u = (u_1, \ldots, u_d) \in (0,1)^d$. The latter satisfies

$$\lim_{n\to\infty} \int_{(0,1)^d} \left[\sqrt{n}\left\{\sqrt{c_{\theta_n}(u)} - 1\right\} - \frac{\delta}{2}\,\dot{c}_{\theta_0}(u)\right]^2 \mathrm{d}u_d \cdots \mathrm{d}u_1 = 0.$$

(ii) The following identity holds for every $u = (u_1, \ldots, u_d) \in (0,1)^d$:

$$\dot{C}_{\theta_0}(u) = \lim_{\theta\to\theta_0} \frac{\partial}{\partial\theta}\, C_\theta(u) = \int_0^{u_1} \cdots \int_0^{u_d} \dot{c}_{\theta_0}(v)\,\mathrm{d}v_d \cdots \mathrm{d}v_1.$$

The following result is a straightforward extension of Proposition 1 of Genest et al. (2005a). Before stating it, let

$$\Lambda(u,u') = C_{\theta_0}(u \wedge u') + (d-1)C_{\theta_0}(u)C_{\theta_0}(u') - C_{\theta_0}(u)C_{\theta_0}(u')\sum_{j=1}^{d}\left(\frac{u_j \wedge u'_j}{u_j u'_j}\right).$$

Proposition 6.1. *Suppose that the underlying copula of a given population belongs to a family $\mathcal{C}$ whose members satisfy assumptions* (i) *and* (ii)*. Then, under Q_n, the empirical process $\mathbb{C}_n = n^{1/2}\left(C_n - C_{\theta_0}\right)$ converges in law in the space $\mathcal{D}([0,1]^d)$ to a continuous Gaussian process $\mathbb{C} + \delta\dot{C}_{\theta_0}$ with covariance structure Λ.*

Proof. Introduce $U_{ij}^{(n)} = F_j\left(X_{ij}^{(n)}\right)$ and define

$$\mathbb{A}_n(u) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\prod_{j=1}^{d}\mathbf{1}\left(U_{ij}^{(n)} \le u_j\right) - \prod_{j=1}^{d} u_j\right\}.$$

Let also

$$\Psi_{j,n}(u) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\left(U_{ij}^{(n)} \le u\right).$$

Since $R_{ij}^{(n)} = n\Psi_{jn}\left(U_{ij}^{(n)}\right)$, one deduces from equations (6.1) and (6.2) that

$$\mathbb{C}_n(u) = \mathbb{A}_n\left\{\Psi_{1,n}^{-1}(u_1),\ldots,\Psi_{d,n}^{-1}(u_d)\right\} + \sqrt{n}\left\{\prod_{j=1}^{d}\Psi_{j,n}^{-1}(u_j) - \prod_{j=1}^{d}u_j\right\}. \quad (6.3)$$

Under Assumption (i), an application of Theorem 3.10.12 of van der Vaart (1996) implies that under Q_n, the sequence $(\mathbb{A}_n)$ of processes converges in $\mathcal{D}\left([0,1]^d\right)$ to a continuous Gaussian limit of the form $\mathbb{A} + \delta\dot{C}_{\theta_0}$, where $\dot{C}_{\theta_0}$ is defined as in Assumption (ii).

As a consequence of this result, one has under Q_n that the univariate process $\mathbb{A}_n(\mathbf{1}, u_j, \mathbf{1}) = \sqrt{n}\,\{\Psi_{j,n}(u_j) - u_j\}$ converges in $\mathcal{D}\,([0,1])$ to

$$\mathbb{A}(\mathbf{1}, u_j, \mathbf{1}) + \delta\dot{C}_{\theta_0}(\mathbf{1}, u_j, \mathbf{1}) = \mathbb{A}(\mathbf{1}, u_j, \mathbf{1})\,,$$

since

$$\dot{C}_{\theta_0}(\mathbf{1}, u_j, \mathbf{1}) = \lim_{\theta\to\theta_0} \frac{C_\theta(\mathbf{1}, u_j, \mathbf{1}) - C_{\theta_0}(\mathbf{1}, u_j, \mathbf{1})}{\theta - \theta_0} = \lim_{\theta\to\theta_0} \frac{u_j - u_j}{\theta - \theta_0} = 0.$$

From identities (11) and (12) in Chapter 3 of Shorack & Wellner (1984), one has for each $1 \le j \le d$ that

$$\sup_{0\le u\le 1} |\Psi_{j,n}(u) - u| = \sup_{0\le u\le 1} \left|\Psi_{j,n}^{-1}(u) - u\right|$$

tends to zero in probability, from which it follows that $\sqrt{n}\left\{\Psi_{j,n}^{-1}(u_j) - u_j\right\}$ converges in $\mathcal{D}\,([0,1])$ to $-\mathbb{A}(\mathbf{1}, u_j, \mathbf{1})$.

Finally, since the second summand in equation (6.3) can be written as

$$\sum_{k=1}^{d} \left(\prod_{i=0}^{k-1} u_i\right) \left(\prod_{j=k+1}^{d} \Psi_{j,n}^{-1}(u_j)\right) \sqrt{n}\left\{\Psi_{k,n}^{-1}(u_k) - u_k\right\},$$

one concludes that under Q_n, the process $\mathbb{C}_n$ converges in $\mathcal{D}\left([0,1]^d\right)$ to $\mathbb{C} + \delta\dot{C}_{\theta_0}$, where

$$\mathbb{C}(u) = \mathbb{A}(u) - \sum_{k=1}^{d} \mathbb{A}(\mathbf{1}, u_k, \mathbf{1}) \prod_{i\neq k} u_i$$

is the limiting process of $\mathbb{C}_n$ under P_n identified by Gänssler & Stute (1987). Finally, by a straightforward computation, one may show that the limiting covariance function of $\mathbb{C}_n$ is Λ, noting that $\delta\dot{C}_{\theta_0}$ is a deterministic term. $\diamondsuit$

Deheuvels (1981a,b,c) proposed to decompose $\mathbb{C}_n$ into a collection of asymptotically independent, centered Gaussian processes having a simple covariance function under the null hypothesis of independence. Specifically, for each $A \subset \mathcal{S}_d$, define the linear operator $\mathcal{M}_A$ such that

$$\begin{aligned}\mathbb{G}_{A,n}(u) = \mathcal{M}_A\{\mathbb{C}_n(u)\} &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \prod_{j \in A} \{\mathbf{1}\left(R_{ij} \leq n u_j\right) - u_j\} \\ &= \sum_{B \subset A} (-1)^{|A \setminus B|} \mathbb{C}_n\left(u^B\right) C_{\theta_0}\left(u^{A \setminus B}\right), \quad (6.4)\end{aligned}$$

where the notation $u^B \in [0,1]^d$ is defined as

$$u_j^B = \begin{cases} u_j & \text{if } j \in B, \\ 1 & \text{if } j \notin B. \end{cases}$$

The result is that under P_n, the empirical process $\mathbb{C}_n$ is decomposed into $2^d - d - 1$ sub-processes that converge jointly to a vector of continuous centered Gaussian processes $\mathbb{G}_A = \mathcal{M}_A(\mathbb{C})$ with covariance structure

$$\Gamma_{A,A'}(u,v) = \operatorname{cov}\left\{\mathbb{G}_A(u), \mathbb{G}_{A'}(v)\right\} = \begin{cases} \prod_{j \in A} \gamma(u_j, v_j) & \text{if } A = A', \\ 0 & \text{if } A \neq A', \end{cases}$$

where $\gamma(u,v) = \min(u,v) - uv$. In other words, $\mathbb{G}_{A,n}$ are asymptotically independent and behave as Brownian bridges.

The next result, which is a straightforward consequence of Proposition 6.1, gives the asymptotic representation of $\mathbb{G}_{A,n}$ under Q_n.

Corollary 6.1. *Let $\mathcal{C}$ be a given family of copulas whose members satisfy assumptions* (i) *and* (ii)*. Then, under Q_n, the empirical processes $\mathbb{G}_{A,n}$ converge jointly in $\mathcal{D}([0,1]^d)$ to a vector of continuous Gaussian processes $\mathbb{G}_A + \delta q_A$ with covariance structure $\Gamma_{A,A'}$, where $q_A = \mathcal{M}_A(\dot{C}_{\theta_0})$.*

Proof. Since $\mathcal{M}_A$ is a continuous, linear operator, $\mathbb{G}_{A,n} = \mathcal{M}_A(\mathbb{C}_n)$ converges in law in $\mathcal{D}\left([0,1]^d\right)$ to $\mathcal{M}_A(\mathbb{C} + \delta\dot{C}_{\theta_0}) = G_A + \delta\mathcal{M}_A(\dot{C}_{\theta_0})$, while the covariance structure follows from the fact that $\delta\mathcal{M}_A(\dot{C}_{\theta_0})$ is a deterministic term. $\diamondsuit$

6.3 Examples

The drift parameter $q_A(u)$ identified in Corollary 6.1 can be computed explicitly in many families of copulas. A few examples are provided below.

The following lemma, which is called the Möbius inversion formula, will be used repeatedly in this section for computing q_A for various families of copulas. Its proof can be found in, e.g., Spitzer (1974), p. 127.

Lemma 6.1. *Let f be a function defined on the subsets B of $\mathcal{S}_d$. For any $A \subset \mathcal{S}_d$, set*

$$F(A) = \sum_{B \subset A} f(B).$$

Then

$$f(A) = \sum_{B \subset A} (-1)^{|A \setminus B|} F(B).$$

6.3.1 The multivariate equicorrelated normal copula

Let N denote the distribution function of a standard normal random variable. The multivariate normal copula with $d \times d$ correlation matrix $\Sigma = (\sigma_{jk})$ such that $\sigma_{jj} = 1$ for $j \in \mathcal{S}_d$ is defined by

$$C_\Sigma(u) = H_\Sigma\left\{N^{-1}(u_1), \ldots, N^{-1}(u_d)\right\},$$

where

$$H_\Sigma(x) = \int_{-\infty}^{x_d} \cdots \int_{-\infty}^{x_1} \frac{1}{(2\pi)^{d/2}\,|\Sigma|^{1/2}} \exp\left(-y^\top \Sigma^{-1} y/2\right) \mathrm{d}y_1 \cdots \mathrm{d}y_d.$$

Consider the equicorrelated case in which $\sigma_{jk} = \rho$ for all $j \neq k$. Write $H_\rho = H_\Sigma$ and $\dot{H}_\rho = \mathrm{d}H_\rho/\,\mathrm{d}\rho$. One can show that

$$\frac{d}{d\rho}\,|\Sigma|^{-1/2}\Big|_{\rho=0} = 0 \quad \text{and} \quad -\frac{1}{2}\,\frac{d}{d\rho} x^\top \Sigma^{-1} x\Big|_{\rho=0} = \sum_{i<j} x_i x_j,$$

where the latter identity follows from the fact that

$$x^\top \Sigma^{-1} x = \frac{(1-\rho)^{d-2}}{|\Sigma|}\left[\{(d-2)\rho + 1\}\sum_{k=1}^d x_k^2 - 2\rho \sum_{i<j} x_i x_j\right].$$

Now if $N'(t) = \mathrm{d}N(t)/\,\mathrm{d}t$, it follows that

$$\begin{aligned}
\lim_{\rho\to 0} \dot{H}_\rho(x) &= \int_{-\infty}^{x_d} \cdots \int_{-\infty}^{x_1} \left(\sum_{i<j} y_i y_j\right)\left\{\prod_{k=1}^d N'(y_k)\right\} \mathrm{d}y_1 \cdots \mathrm{d}y_d \\
&= \sum_{i<j} \int_{-\infty}^{x_d} \cdots \int_{-\infty}^{x_1} y_i y_j \left\{\prod_{k=1}^d N'(y_k)\right\} \mathrm{d}y_1 \cdots \mathrm{d}y_d \\
&= \sum_{i<j} \left\{\prod_{k\neq i,j} N(x_k)\right\} \int_{-\infty}^{x_i} y_i N'(y_i)\,\mathrm{d}y_i \int_{-\infty}^{x_j} y_j N'(y_j)\,\mathrm{d}y_j \\
&= \sum_{i<j} \left\{\prod_{k\neq i,j} N(x_k)\right\} N'(x_i) N'(x_j) \\
&= H_0(x) \sum_{i<j} \frac{N'(x_i)}{N(x_i)} \frac{N'(x_j)}{N(x_j)}.
\end{aligned}$$

Thus,

$$\begin{aligned}\dot{C}_{\theta_0}(u) &= \lim_{\rho\to 0} \dot{H}_\rho \left\{N^{-1}(u_1), \ldots, N^{-1}(u_d)\right\} \\ &= C_{\theta_0}(u) \sum_{i<j} \frac{N'\left\{N^{-1}(u_i)\right\}}{u_i} \frac{N'\left\{N^{-1}(u_j)\right\}}{u_j}.\end{aligned}$$

An application of the Möbius inversion formula, with

$$f(A) = \mathbf{1}(|A| = 2) \prod_{j\in A} \frac{N'\left\{N^{-1}(u_j)\right\}}{u_j},$$

then yields

$$q_A(u) = \mathbf{1}(|A| = 2) \prod_{j\in A} N'\left\{N^{-1}(u_j)\right\}.$$

Hence, when looking at the Möbius decomposition of the multivariate normal model, only the $G_{A,n}$ with $|A| = 2$ have a limiting distribution that differs under the null hypothesis and the contiguous sequence of alternative hypotheses. Accordingly, tests of independence should only be based on the latter; the inclusion of functions of $G_{A,n}$ for any $|A| > 2$ would be useless, as they would contribute nothing to the overall power of the procedure. This observation does not come as a total surprise, since the multivariate Gaussian dependence structure is completely characterized by the pairwise interactions among the variables.

6.3.2 One-parameter multivariate Farlie–Gumbel–Morgenstern copula

Another example is the multivariate extension of the Farlie–Gumbel–Morgenstern copula family of distributions defined for $\theta \in [-1, 1]$ by

$$C_\theta(u) = C_{\theta_0}(u) + \theta \prod_{j=1}^{d} u_j(1-u_j), \quad u \in [0,1]^d.$$

The value $\theta_0 = 0$ corresponds to independence. It follows easily that

$$\dot{C}_{\theta_0}(u) = \prod_{j=1}^{d} u_j(1-u_j).$$

Since $\dot{C}_{\theta_0}(u^B)$ vanishes whenever $B \neq \mathcal{S}_d$, one finds that

$$q_A(u) = (-1)^d \dot{C}_{\theta_0}(u)\, \mathbf{1}\,(A = \mathcal{S}_d)\,.$$

Thus, in this case, q_A vanishes unless $A = \mathcal{S}_d$, which implies that in contrast to the equicorrelated multivariate normal case, tests of independence based on functions $G_{A,n}$ with any $|A| < d$ would have no power in the neighborhood of independence.

6.3.3 Archimedean copulas

A copula is called Archimedean (Genest and MacKay 1986; Nelsen 1999, Chapter 4) whenever it can be expressed in the form

$$C(u) = \phi^{-1}\left\{\phi(u_1) + \cdots + \phi(u_d)\right\}$$

in terms of a univariate generator $\phi : (0,1] \to [0,\infty)$ that satisfies $\phi(1) = 0$ and

$$(-1)^j \frac{\mathrm{d}^j}{\mathrm{d}t^j} \phi^{-1}(t) > 0$$

for every $j \in \mathcal{S}_d$.

Proposition 6.2. *Let $\mathcal{C}$ be a parametric family of Archimedean copulas with generator ϕ_θ satisfying*

$$\lim_{\theta \to \theta_0} \phi_\theta(t) = -\log t \quad \textit{and} \quad \lim_{\theta \to \theta_0} \phi'_\theta(t) = -\frac{1}{t} \,.$$

Further assume that $\mathcal{C}$ meets conditions (i) *and* (ii)*. Then*

$$\frac{\dot{C}_{\theta_0}(u)}{C_{\theta_0}(u)} = \dot{\phi}_{\theta_0}\left\{C_{\theta_0}(u)\right\} - \sum_{j=1}^{d} \dot{\phi}_{\theta_0}(u_j) \tag{6.5}$$

and

$$q_A(u) = C_{\theta_0}\left(u^A\right) \sum_{B \subset A} (-1)^{|A \setminus B|} \dot{\phi}_{\theta_0}\left\{C_{\theta_0}\left(u^B\right)\right\}. \tag{6.6}$$

Proof. Noting that $\phi_\theta\left\{C_\theta(u)\right\} = \phi_\theta(u_1) + \cdots + \phi_\theta(u_d)$ and applying the chain rule, one finds

$$\dot{\phi}_\theta\left\{C_\theta(u)\right\} + \dot{C}_\theta(u)\phi'_\theta\left\{C_\theta(u)\right\} = \sum_{j=1}^{d} \dot{\phi}_\theta(u_j).$$

Equation (6.5) follows by taking the limit as $\theta \to \theta_0$ while (6.6) emerges by a straight substitution of $\dot{C}_{\theta_0}$ into the formula for $\mathcal{M}_A$, combined with the fact that when $|A| \geq 2$,

$$\sum_{B \subset A} \sum_{j \in B} (-1)^{|A \setminus B|} \dot{\phi}_{\theta_0}(u_j) = \mathbf{1}(|A| = 1) \prod_{j \in A} \dot{\phi}_{\theta_0}(u_j) = 0$$

because of the Möbius inversion formula. ♢

Example 6.1. *Assumptions* (i) *and* (ii) *can easily be verified for Frank's family of* $d-$*variate copulas (Frank 1979; Nelsen 1986; Genest 1987), whose generator is given by*

$$\phi_\theta(t) = -\log\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right), \quad t \in (0,1]$$

with $\theta \in [\ell_d, \infty)$, *where* $-\infty = \ell_2 < \ell_3 < \cdots < \ell_\infty = 0$. *Here,* $\theta_0 = 0$ *corresponds to independence and* $\dot\phi_{\theta_0}(t) = (t-1)/2$. *In view of Proposition 6.2 and by an application of the Möbius inversion formula, one finds*

$$\dot C_{\theta_0}(u) = \frac{1}{2}\, C_{\theta_0}(u)\left\{d-1+C_{\theta_0}(u) - \sum_{j=1}^{d} u_j\right\}$$

and

$$\begin{aligned} q_A(u) &= \frac{1}{2}\, C_{\theta_0}(u^A) \sum_{B\subset A} (-1)^{|A\setminus B|}\left\{C_{\theta_0}(u^B)-1\right\} \\ &= \frac{1}{2}\, C_{\theta_0}(u^A) \sum_{B\subset A} (-1)^{|A\setminus B|} C_{\theta_0}(u^B) \\ &= \frac{1}{2}\prod_{j\in A}\left(u_j^2 - u_j\right). \end{aligned}$$

Here, $q_A \neq 0$ *for every* $A \subset \mathcal{S}_d$ *with* $|A| \geq 2$.

Remark 6.3. *The* $d-$*variate Archimedean system of Ali et al. (1978) with* $0 \leq \theta \leq 1$ *is generated by*

$$\phi_\theta(t) = \frac{1}{1-\theta}\log\left(\frac{1-\theta}{t}+\theta\right), \quad t \in (0,1]$$

and $\theta_0 = 0$ *corresponds to independence. One gets* $\dot\phi_{\theta_0}(t) = t - 1 - \log t$ *and hence* $q_A(u) = \prod_{j\in A}(u_j^2 - u_j)$, *which is the same, up to a multiplicative constant, than that of Frank's family.*

Also, for another version of the multivariate Farlie–Gumbel–Morgenstern copula, namely

$$C_\theta(u) = C_{\theta_0}(u) + \theta C_{\theta_0}(u) \left\{ d - 1 + C_{\theta_0}(u) - \sum_{j=1}^{d} u_j \right\},$$

one finds that $\dot{C}_{\theta_0}$ *and* q_A *are the same as for Frank's family, up to a multiplicative constant.*

Example 6.2. *Assumptions* (i) *and* (ii) *can also easily be verified for Clayton's* $d-$*variate family of copulas (Clayton 1978), whose generator, defined for all* $\theta \in [0, \infty)$*, is given by*

$$\phi_\theta(t) = \frac{t^{-\theta} - 1}{\theta}, \quad t \in (0, 1].$$

Note that independence corresponds to the value $\theta_0 = 0$*. One finds easily that* $\dot{\phi}_{\theta_0}(t) = (\log t)^2/2$*. It follows from Proposition 6.2 that*

$$\begin{aligned} \dot{C}_{\theta_0}(u) &= \frac{1}{2} C_{\theta_0}(u) \left\{ \left(\sum_{j=1}^{d} \log u_j \right)^2 - \sum_{j=1}^{d} (\log u_j)^2 \right\} \\ &= C_{\theta_0}(u) \sum_{j<k} \log u_j \log u_k, \end{aligned}$$

and

$$q_A(u) = \mathbf{1}(|A| = 2) \prod_{j \in A} u_j \log u_j,$$

the latter identity following from the use of the Möbius inversion formula.

Remark 6.4. *The dependence structure induced by Clayton's copula is also known as the gamma frailty model in survival analysis. It may come somewhat as a surprise that for this model,* $q_A(u) = 0$ *unless* $|A| = 2$*. In other* *words, Clayton's copula shares with the multivariate normal model the property that tests of independence based on terms* $G_{A,n}$ *of the Möbius decomposition with* $|A| > 2$ *would have no power whatsoever in the neighborhood of independence.*

The Gumbel–Barnett system of copulas provides another example of this curious phenomenon. Copulas in this Archimedean class are generated by

$$\phi_\theta(t) = \log(1 - \theta \log t)/\theta, \quad t \in (0, 1]$$

with $\theta \in [0, 1]$ *and* $\theta_0 = 0$ *corresponding to independence. A simple calculation shows that* $\dot{\phi}_{\theta_0}(t) = -(\log t)^2/2$ *and hence*

$$\dot{C}_{\theta_0}(u) = -C_{\theta_0}(u) \sum_{j<k} \log u_j \log u_k,$$

and

$$q_A(u) = -\mathbf{1}(|A| = 2) \prod_{j \in A} u_j \log u_j.$$

These formulas are the same as for Clayton's copula, up to a change in sign.

Example 6.3. *Assumptions* (i) *and* (ii) *can also easily be verified for Gumbel–Hougaard's family of copulas (Gumbel 1960), whose generator is defined for* $\theta \in [0, 1)$ *by*

$$\phi_\theta(t) = |\log t|^{1/(1-\theta)}, \quad t \in (0, 1].$$

In that case, one finds

$$\dot{\phi}_{\theta_0}(t) = -(\log t) \log(\log 1/t),$$

so by Proposition 6.2,

$$\dot{C}_{\theta_0}(u) = C_{\theta_0}(u) \left\{ -\sum_{j=1}^{d} \log u_j \log \left(\sum_{k=1}^{d} \log u_k / \log u_j \right) \right\}$$

and

$$q_A(u) = -C_{\theta_0}\left(u^A\right) \sum_{B \subset A} (-1)^{|A \setminus B|} \left(\sum_{j \in B} \log u_j \right) \log \left(- \sum_{j \in B} \log u_j \right),$$

for $u \in (0,1)^d$. *Here again,* $q_A \neq 0$ *for every* $A \subset \mathcal{S}_d$ *with* $|A| \geq 2$.

6.4 Limiting distributions of Cramér–von Mises functionals

In the absence of information about the marginal distributions of a multivariate population, a valid testing procedure for multivariate independence should be based on some version of the empirical copula process $\mathbb{C}_n$. To improve convergence and reduce bias in finite samples, a centered version of $\mathbb{C}_n$ will be used in the sequel. The latter is defined by

$$\tilde{\mathbb{C}}_n(u) = \sqrt{n} \left\{ C_n(u) - \prod_{j=1}^{d} U_n(u_j) \right\},$$

where U_n is the distribution function of a uniformly distributed random variable on the set $\{1/n, \ldots, n/n\}$. In other words, $U_n(u) = \lfloor nu \rfloor / n$, with $\lfloor x \rfloor$ standing for the integer part of x. It is clear that $\tilde{\mathbb{C}}_n$ and $\mathbb{C}_n$ have the same limiting behavior, so that the asymptotic results of Section 6.2 also apply to $\tilde{C}_n$, and henceforth to $\tilde{G}_{A,n} = \mathcal{M}_A(\tilde{\mathbb{C}}_n)$.

A natural way to test for multivariate independence is to consider a global measure of discrepancy computed from $\tilde{\mathbb{C}}_n$ or from a combination of distances computed for each of the $\tilde{\mathbb{G}}_{A,n}$ taken individually. Two candidates are the

Kolmogorov–Smirnov statistics,

$$S_n = \sup_{u\in[0,1]^d} \left|\tilde{\mathbb{C}}_n(u)\right| \quad \text{and} \quad S_{A,n} = \sup_{u\in[0,1]^d} \left|\tilde{\mathbb{G}}_{A,n}(u)\right|,$$

and the Cramér–von Mises functionals,

$$B_n = \int_{[0,1]^d} \left\{\tilde{\mathbb{C}}_n(u)\right\}^2 \mathrm{d}u \quad \text{and} \quad B_{A,n} = \int_{[0,1]^d} \left\{\tilde{\mathbb{G}}_{A,n}(u)\right\}^2 \mathrm{d}u\,.$$

In the sequel, however, attention will be limited to B_n and $B_{A,n}$. Note that these statistics can be expressed as functions of the ranks through

$$\begin{aligned} B_n &= \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n\prod_{k=1}^d\left(1-\frac{R_{ik}\vee R_{jk}}{n}\right)+n\left\{\frac{(n-1)(2n-1)}{6n^2}\right\}^d \\ &\quad -2\sum_{i=1}^n\prod_{k=1}^d\left\{\frac{n(n-1)-R_{jk}(R_{jk}-1)}{2n^2}\right\} \end{aligned}$$

and

$$B_{A,n} = \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n\prod_{k\in A} D_n(R_{ik},R_{jk}), \tag{6.7}$$

where

$$D_n(s,t) = \frac{(n+1)(2n+1)}{6n^2} + \frac{s(s-1)}{2n^2} + \frac{t(t-1)}{2n^2} - \frac{\max(s,t)}{n}\,.$$

6.4.1 Asymptotics for B_n

Since even under P_n, no explicit Karhunen–Loève expansion for $\mathbb{C}$ is available when $d > 2$, the asymptotic null distribution of B_n cannot be computed analytically. This is due to the unwieldy form of the covariance function Λ.

Now if $\mathcal{C}$ is a family of copulas whose members satisfy assumptions (i) and (ii), one has from Proposition 6.1 that B_n converges in law, under Q_n, to

$$\mathbb{B} = \int_{[0,1]^d} \left\{ \mathbb{C}(u) + \delta \dot{C}_{\theta_0}(u) \right\}^2 \mathrm{d}u, \tag{6.8}$$

which is a Cramér–von Mises functional of a Gaussian process with mean $\delta \dot{C}_{\theta_0}$ and covariance function Λ. In order to approximate the distribution of $\mathbb{B}$, a procedure due to Deheuvels & Martynov (1996) will be adopted. Specifically, the proposed approximation is

$$\tilde{\mathbb{B}} = \frac{1}{m} \|\xi\|^2 = \frac{1}{m} \sum_{i=1}^{m} \xi_i^2, \tag{6.9}$$

where $\xi = \mu(U) + V(U)Z$ is an $m-$variate vector constructed from two independent vectors U and Z whose components are mutually independent and $\mathcal{U}(0,1)$ and $\mathcal{N}(0,1)$, respectively. Furthermore,

$$\mu(u_1, \ldots, u_m) = \delta \left(\dot{C}_{\theta_0}(u_1), \ldots, \dot{C}_{\theta_0}(u_m) \right)^\top$$

and $V(u_1, \ldots, u_m)$ is the Cholesky decomposition of the covariance matrix Σ with components

$$\Sigma_{jk} = \Lambda(u_j, u_k), \quad 1 \leq j, k \leq m.$$

Deheuvels & Martynov (1996) show that the variance of the approximation error, $\mathrm{var}(\tilde{\mathbb{B}} - \mathbb{B})$, is $O(1/m)$.

6.4.2 Asymptotics for $B_{A,n}$

In view of the familiar form of the covariance structure of $\mathbb{G}_A$, which is in fact a product of covariance functions associated to Brownian bridges, it follows from standard theory (see, e.g., Shorack and Wellner 1984, p. 213) that the limiting process $\mathbb{G}_A$ admits, under P_n, the representation

$$\mathbb{G}_A(u) = \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma^{1/2} Z_\gamma f_\gamma(u), \quad \gamma = (\gamma_j)_{j \in A}, \tag{6.10}$$

where the Z_γ are independent $\mathcal{N}(0,1)$ random variables and

$$\lambda_\gamma = \prod_{j \in A} (\pi \gamma_j)^{-2}, \quad f_\gamma(u) = \prod_{j \in A} \sqrt{2} \sin(\gamma_j \pi u_j), \quad 0 \leq u_j \leq 1.$$

An idea first proposed by Deheuvels (1981a,b,c) and latter exploited by Genest & Rémillard (2004) is to base a test for multivariate independence on some combination of the asymptotically independent statistics $B_{A,n}$. In view of Corollary 6.1, the limiting distribution of $B_{A,n}$ under Q_n, is given by

$$\mathbb{B}_A = \int_{(0,1)^d} \{\mathbb{G}_A(u) + \delta q_A(u)\}^2 \, \mathrm{d}u.$$

As already shown by Deheuvels (1981a,b,c), formula (6.10) implies that

$$\mathbb{Q}_A = \int_{(0,1)^d} \{\mathbb{G}_A(u)\}^2 \, \mathrm{d}u = \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma Z_\gamma^2, \quad \lambda_\gamma = \prod_{j \in A} \lambda_{\gamma_j}.$$

It follows that

$$\mathbb{B}_A = \mathbb{Q}_A + 2\delta \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma I_{\gamma,A} Z_\gamma + \delta^2 I_A,$$

where

$$I_A = \int_{[0,1]^d} \{q_A(u)\}^2 \, \mathrm{d}u \quad \text{and} \quad I_{\gamma,A} = \lambda_\gamma^{-1/2} \int_{(0,1)^d} q_A(u) f_\gamma(u) \, \mathrm{d}u.$$

The limiting distribution of $B_{A,n}$ under Q_n is given below.

Proposition 6.3. *If $\mathcal{C}$ is a family of copulas whose members satisfy assumptions* (i) *and* (ii)*, the asymptotic distribution of $B_{A,n}$, under Q_n, is given by the weighted sum*

$$\mathbb{B}_A = \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma \left(Z_\gamma + \delta I_{\gamma,A} \right)^2 .$$

Proof. Parseval's identity states that

$$I_A = \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma I_{\gamma,A}^2.$$

Making this substitution into the representation of $\mathbb{B}_A$ and exploiting the representation of $\mathbb{Q}_A$, one gets

$$\begin{aligned}
\mathbb{B}_A &= \mathbb{Q}_A + 2\delta \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma I_{\gamma,A} Z_\gamma + \delta^2 I_A, \\
&= \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma Z_\gamma^2 + 2\delta \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma I_{\gamma,A} Z_\gamma + \delta^2 \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma I_{\gamma,A}^2 \\
&= \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma \left(Z_\gamma + \delta I_{\gamma,A}\right)^2.
\end{aligned}$$

Hence the result. ◊

Note there also exists a representation like (6.10) for $\mathbb{C}$, but the weighted factors are unknown. Nevertheless, their sum is finite. Using the same technique as in the proof of Proposition 6.3 and using representation (6.8), it follows that

$$\mathbb{B} = \sum_{\gamma \in \mathbb{N}^d} \tilde{\lambda}_\gamma (\tilde{Z}_\gamma + \delta \tilde{I}_\gamma)^2, \tag{6.11}$$

where

$$\sum_{\gamma \in \mathbb{N}^d} \tilde{\lambda}_\gamma = \int_{[0,1]^d} \int_{[0,1]^d} \Gamma(u,v)\, \mathrm{d}v\, \mathrm{d}u < \infty$$

and

$$\sum_{\gamma \in \mathbb{N}^d} \tilde{\lambda}_\gamma \tilde{I}_\gamma^2 = \int_{[0,1]^d} \left\{\dot{C}_{\theta_0}(u)\right\}^2 \mathrm{d}u < \infty.$$

The unknown quantities $\tilde{\lambda}_\gamma$ and $\tilde{I}_\gamma$ could be approximated numerically.

For many systems of distributions, the drift function has the simple form $q_A(u) = \prod_{j\in A} q(u_j)$. In that case, it is easy to see that

$$I_{\gamma,A} = 2^{|A|/2}\lambda_\gamma^{-1/2}\prod_{j\in A} f(\gamma_j), \quad \text{where} \quad f(k) = \int_0^1 q(u)\sin(k\pi u)\ \mathrm{d}u. \tag{6.12}$$

Example 6.4 (*Equicorrelated Gaussian copulas*). *For this model, it follows from* (6.12) *and results from Section 6.3.1 that*

$$I_{\gamma,A} = 2\pi^2\left(\prod_{j\in A}\gamma_j\, g(\gamma_j)\right)\mathbf{1}(|A|=2),$$

where

$$g(k) = \int_0^1 N'\left\{N^{-1}(u)\right\}\sin(k\pi u)\ \mathrm{d}u = \int_{\mathbb{R}}\left\{N'(t)\right\}^2\sin\left\{k\pi N(t)\right\}\ \mathrm{d}t.$$

Example 6.5 (*Farlie–Gumbel–Morgenstern copulas*). *For this model, results from Section 6.3.2 imply that for any* $\gamma\in\mathbb{N}^{|A|}$,

$$I_{\gamma,A} = \begin{cases} (-1)^d 2^{5d/2}\lambda_\gamma\mathbf{1}\left(A=\mathcal{S}_d\right), & \text{if } \gamma_1,\ldots,\gamma_d \text{ are all odd;}\\ 0, & \text{otherwise.}\end{cases}$$

Example 6.6 (*Frank and Ali–Mikhail–Haq copulas*). *For these models, calculations based on material from Example 6.1 lead to*

$$I_{\gamma,A} = \begin{cases} (-1)^{|A|}2^{5|A|/2-1}\lambda_\gamma, & \text{if } \gamma_j, j\in A \text{ are all odd;}\\ 0, & \text{otherwise.}\end{cases}$$

Example 6.7 (*Clayton and Gumbel–Barnett distributions*). *For these models, the observations already made in Example 6.2 and Remark 6.4 yield*

$$I_{\gamma,A} = \frac{2}{\pi^2}\left(\prod_{j\in A}\frac{SI(\gamma_j\pi)}{\gamma_j}\right)\mathbf{1}(|A|=2),$$

where $SI(x) = \int_0^x t^{-1}\sin(t)\ \mathrm{d}t$.

6.4.3 Combination of independent statistics

A simple test of independence consists in rejecting the null hypothesis whenever the observed value of B_n exceeds the $(1-\alpha)-$percentile of its asymptotic distribution under P_n. The critical value $q_B(\alpha)$ such that

$$\lim_{n\to\infty} \mathrm{P}\{B_n > q_B(\alpha) \mid P_n\} = \alpha$$

is easily approximated using formula (6.9) with $\delta = 0$. These values, for $\alpha = 0.05$, are to be found in Table 6.1.

However, potentially more efficient methods could be based on some combination of the asymptotically independent $B_{A,n}$ with $A \in \mathcal{S}_d$. Here, four such procedures are considered. The first is inspired from Ghoudi et al. (2001), while the second and third were investigated by Genest & Rémillard (2004).

(1) *Linear combination rule.* Base the test on $L_n = \sum_{|A|>1} B_{A,n}$. Table 6.1 gives approximations to the values of $q_L(\alpha)$ such that $\mathrm{P}\{L_n > q_L(\alpha) \mid P_n\} = \alpha$ as $n \to \infty$.

(2) *Dependogram method.* Base the test on

$$M_n = \max_{|A|>1} \left\{ \frac{B_{A,n}}{q_{|A|}(\alpha')} \right\},$$

where $q_{|A|}(\alpha)$ is such that $\lim_{n\to\infty} \mathrm{P}\{B_{A,n} > q_{|A|}(\alpha') \mid P_n\} = 1-\alpha$ and

$$\alpha' = 1 - (1-\alpha)^{1/(2^d-d-1)} \tag{6.13}$$

is chosen so that by performing each test at level α', the global level of the procedure is α. The first few simulated critical values $q_{|A|}$ are reported in Table 6.1. Note that $\mathrm{P}(M_n > 1|P_n) \to \alpha$ as $n \to \infty$.

(3) *Fisher's approach.* Base the test on

$$T_n = -2 \sum_{|A|>1} \log \left\{ 1 - F_{A,n}(B_{A,n}) \right\},$$

where $F_{A,n}(x) = \mathrm{P}\left(B_{A,n} \le x\right)$. Under P_n, T_n converges in distribution to

$$-2 \sum_{|A|>1} \log \left\{ 1 - q_{|A|}^{-1}(\mathbb{B}_A) \right\},$$

which is chi-square with $2(2^d - d - 1)$ degrees of freedom. Hence, the critical value of a test based on T_n is given by

$$q_T(\alpha) = K^{-1}(1-\alpha), \quad \text{where } K(x) = \mathrm{P}\left(\chi^2_{2(2^d-d-1)} \le x\right).$$

Table 6.1 reports the asymptotic values of $q_T(0.05)$ for $d = 3, 4, 5$, as per Genest & Rémillard (2004).

Finally, the following alternative weighing of the individual tests is inspired from Proposition A.1 in the Appendix. The latter implies that whenever $B_{A,n}$ is large,

$$-2 \log \left\{ 1 - F_{A,n}(B_{A,n}) \right\} \approx \pi^{2|A|} B_{A,n}.$$

Therefore,

$$W_n = \sum_{|A|>1} \pi^{2|A|} B_{A,n}$$

is an approximation for T_n, which should be good under fixed alternatives C_θ, since in that case $B_{A,n}/n$ tends to a positive constant. (It is not necessarily precise under all sequence of contiguous alternatives C_{θ_n}, however.)

(4) *Weighted linear combination rule.* Base the test on W_n. Table 6.1 gives approximations to the values of $q_W(\alpha)$ such that $\mathrm{P}\{W_n > q_L(\alpha) \mid P_n\} = \alpha$ as $n \to \infty$.

Note in passing that Littell & Folks (1973) showed that, based on Bahadur's relative efficiency, Fisher's method is optimal, among a large class of combining procedures. This is not to say, however, that procedure (3) is optimal relative to local power, or with respect to any other efficiency measure based thereon. For this reason, comparisons based on local power functions are considered next.

6.5 Comparison of local power functions

One way to compare competing test procedures for multivariate independence is to evaluate their asymptotic power function in a neighborhood of $\theta = \theta_0$, i.e., under copula alternatives C_{θ_n} that form a contiguous sequence (Q_n). Specifically, let S_n be some statistic for independence with asymptotic critical value $q_S(\alpha)$. The associated local power function is defined as

$$\beta_S(\alpha, \delta) = \lim_{n\to\infty} \mathrm{P}\left\{S_n > q_S(\alpha) \;\mid Q_n\right\}.$$

Analytic expressions for this function are given in Subsection 6.5.1 for the combined test statistics L_n, M_n, and W_n, respectively. No similar form can be obtained for B_n or for Fisher's procedure T_n; however, see Remark 6.5. For numerical comparisons, see Subsections 6.5.2 and 6.5.3.

6.5.1 Analytic expressions for local power functions

A result that will prove useful in the sequel is the formula of Gil-Pelaez (1951). It says that if X is a random variable with characteristic function ϕ_X, then

$$\mathrm{P}(X > x) = \frac{1}{2} + \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{Im}\left\{t^{-1} e^{-ixt}\phi_X(t)\right\}\, \mathrm{d}t,$$

where $\mathrm{Im}(z)$ denotes the imaginary part of any complex number z.

To use this identity in the present context, let

$$\begin{aligned}\hat{\eta}(t,\delta) &= (1-2it)^{-1/2}\exp\left(\frac{i\delta^2 t}{1-2it}\right)\\ &= (1+4t^2)^{-1/4}\, e^{-2t^2\delta^2/(1+4t^2)}\, e^{i\arctan(2t)/2 + it\delta^2/(1+4t^2)}\end{aligned}$$

be the characteristic function of a non-central chi-square variable $(Z+\delta)^2$, where $Z \sim \mathcal{N}(0,1)$. From Proposition 6.3, it follows that

$$\begin{aligned}\phi_{\mathbb{B}_A}(t,\delta) = \mathrm{E}\left(e^{it\mathbb{B}_A}\right) &= \prod_{\gamma\in\mathbb{N}^{|A|}} \hat{\eta}\left(\lambda_\gamma t,\, \delta I_{\gamma,A}\right)\\ &= \xi_A(t)\, e^{-2\delta^2 t^2 \kappa_{A,1}(t)}\, e^{i\kappa_{A,2}(t) + i\delta^2\kappa_{A,3}(t)},\end{aligned}$$

where

$$\begin{aligned}\xi_A(t) &= \prod_{\gamma\in\mathbb{N}^{|A|}} (1+4t^2\lambda_\gamma^2)^{-1/4}, \quad \kappa_{A,1}(t) = \sum_{\gamma\in\mathbb{N}^{|A|}} \lambda_\gamma^2 I_{\gamma,A}^2/(1+4t^2\lambda_\gamma^2),\\ \kappa_{A,2}(t) &= \frac{1}{2}\sum_{\gamma\in\mathbb{N}^{|A|}} \arctan(2t\lambda_\gamma), \quad \kappa_{A,3}(t) = t\sum_{\gamma\in\mathbb{N}^{|A|}} \lambda_\gamma I_{\gamma,A}^2/(1+4t^2\lambda_\gamma^2).\end{aligned}$$

Calling again on Proposition 6.3, one can see that the limiting distribution of L_n under Q_n is

$$L = \sum_{|A|>1} \mathbb{B}_A,$$

so that its asymptotic characteristic function is given by

$$\phi_L(t,\delta) = \prod_{|A|>1} \phi_{\mathbb{B}_A}(t,\delta) = \xi(t)\, e^{-2\delta^2 t^2 \kappa_1(t)}\, e^{i\kappa_2(t)+i\delta^2\kappa_3(t)},$$

where

$$\xi(t) = \prod_{|A|>1} \xi_A(t) \quad \text{and} \quad \kappa_i(t) = \sum_{|A|>1} \kappa_{A,i}(t), \quad i=1,2,3.$$

An application of the Gil–Pelaez formula then yields the following result.

Proposition 6.4. *If $\mathcal{C}$ is a family of copulas whose members satisfy assumptions* (i) *and* (ii), *then under Q_n, one has $\beta_L(\alpha,\delta) = \mathrm{P}\{L > q_L(\alpha)\}$, where*

$$\mathrm{P}(L > x) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\sin\{\kappa(x,t)\}}{t\zeta(t)}\, \mathrm{d}t,$$

with

$$\kappa(x,t) = -\frac{xt}{2} + \frac{1}{2}\sum_{|A|>1}\sum_{\gamma\in\mathbb{N}^{|A|}} \left\{\arctan(\lambda_\gamma t) + \delta^2 \frac{\lambda_\gamma I_{\gamma,A}^2 t}{1+\lambda_\gamma^2 t^2}\right\}$$

and

$$\zeta(t) = \exp\left(\frac{\delta^2 t^2}{2}\sum_{|A|>1}\sum_{\gamma\in\mathbb{N}^{|A|}} \frac{\lambda_\gamma^2 I_{\gamma,A}^2}{1+\lambda_\gamma^2 t^2}\right) \prod_{|A|>1}\prod_{\gamma\in\mathbb{N}^{|A|}} \left(1+\lambda_\gamma^2 t^2\right)^{1/4}.$$

Next,

$$\begin{aligned}\lim_{n\to\infty} \mathrm{P}(M_n > 1) &= 1 - \prod_{|A|>1} \mathrm{P}\left\{\mathbb{B}_A \le q_{|A|}(\alpha')\right\} \\ &= 1 - \prod_{|A|>1} \left\{1 - \beta_A(\alpha',\delta)\right\}. \qquad (6.14)\end{aligned}$$

The following result is a consequence of this observation.

Proposition 6.5. *Let M be the limit in distribution of M_n under Q_n, and define β_M to be its associated local power function. If $\mathcal{C}$ is a family of copulas whose members satisfy assumptions* (i) *and* (ii)*, then under Q_n, $\beta_M(\alpha, \delta)$ is given by* (6.14) *and for all $x > 0$,*

$$\mathrm{P}(\mathbb{B}_A > x) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\sin\{\kappa_A(x,t)\}}{t\zeta_A(t)} \, \mathrm{d}t,$$

with

$$\kappa_A(x,t) = -\frac{xt}{2} + \frac{1}{2} \sum_{\gamma \in \mathbb{N}^{|A|}} \left\{ \arctan(\lambda_\gamma \, t) + \delta^2 \, \frac{\lambda_\gamma I^2_{\gamma,A} t}{1 + \lambda^2_\gamma \, t^2} \right\}$$

and

$$\zeta_A(t) = \exp\left(\frac{\delta^2 t^2}{2} \sum_{\gamma \in \mathbb{N}^{|A|}} \frac{\lambda^2_\gamma I^2_{\gamma,A}}{1 + \lambda^2_\gamma \, t^2} \right) \prod_{\gamma \in \mathbb{N}^{|A|}} \left(1 + \lambda^2_\gamma t^2\right)^{1/4}.$$

Next, one can see that the limiting distribution of W_n under Q_n is

$$W = \sum_{|A|>1} \pi^{2|A|} \mathbb{B}_A,$$

so that its asymptotic characteristic function is given by

$$\phi_W(t,\delta) = \prod_{|A|>1} \phi_{\mathbb{B}_A}\left(t\pi^{2|A|}, \delta\right) = \xi_W(t) \, e^{-2\delta^2 t^2 \pi^{4|A|} \kappa_{1,W}(t)} \, e^{i\kappa_{2,W}(t) + i\delta^2 \kappa_{3,W}(t)},$$

where

$$\xi_W(t) = \prod_{|A|>1} \xi_A\left(t\pi^{2|A|}\right) \quad \text{and} \quad \kappa_{i,W}(t) = \sum_{|A|>1} \kappa_{A,i}\left(t\pi^{2|A|}\right), \quad i = 1,2,3.$$

The following proposition gives the local power function of W.

Proposition 6.6. *If $\mathcal{C}$ is a family of copulas whose members satisfy assumptions* (i) *and* (ii)*, then under Q_n, one has $\beta_W(\alpha, \delta) = \mathrm{P}\{W > q_W(\alpha)\}$, where*

$$\mathrm{P}(W > x) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\sin\{\kappa_W(x,t)\}}{t\zeta_W(t)} \, \mathrm{d}t,$$

with

$$\kappa_W(x,t) = -\frac{xt}{2} + \frac{1}{2}\sum_{|A|>1}\sum_{\gamma\in\mathbb{N}^{|A|}}\left\{\arctan(\lambda_\gamma\,\pi^{2|A|}t) + \delta^2\,\frac{\lambda_\gamma\pi^{2|A|}I_{\gamma,A}^2 t}{1+\lambda_\gamma^2\pi^{4|A|}\,t^2}\right\}$$

and

$$\zeta(t) = \exp\left(\frac{\delta^2 t^2}{2}\sum_{|A|>1}\sum_{\gamma\in\mathbb{N}^{|A|}}\frac{\lambda_\gamma^2\pi^{4|A|}I_{\gamma,A}^2}{1+\lambda_\gamma^2\pi^{4|A|}\,t^2}\right)\prod_{|A|>1}\prod_{\gamma\in\mathbb{N}^{|A|}}\left(1+\lambda_\gamma^2\pi^{4|A|}t^2\right)^{1/4}.$$

Remark 6.5. *For Fisher's test, no explicit representation for $\beta_T(\alpha,\delta)$ seems possible. However, under Q_n, T_n converges in distribution to*

$$-2\sum_{|A|>1}\log\left\{1 - q_{|A|}^{-1}\left(\mathbb{B}_A\right)\right\},$$

where

$$q_{|A|}^{-1}(x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \frac{\sin\{\kappa_0(x,t)\}}{t\zeta_0(t)}\,\mathrm{d}t,$$

with

$$\kappa_0(x,t) = -\frac{xt}{2} + \frac{1}{2}\sum_{\gamma\in\mathbb{N}^{|A|}}\arctan(\lambda_\gamma t) \quad \textit{and} \quad \zeta_0(t) = \prod_{\gamma\in\mathbb{N}^{|A|}}\left(1+\lambda_\gamma^2 t^2\right)^{1/4}.$$

An approximation is then given by

$$\hat\beta_T(\alpha,\delta) = \frac{1}{N}\sum_{i=1}^N \mathbf{1}\left\{T_i > q_T(\alpha)\right\},$$

where

$$T_i = -2\sum_{|A|>1}\log\left\{1 - q_{|A|}^{-1}\left(\mathbb{B}_{A,i}\right)\right\}$$

and $\mathbb{B}_{A,i}$ is a truncated version of a random variable whose distribution is the same as that of $\mathbb{B}_A$.

6.5.2 Power comparisons for statistics using all $A \subset \mathcal{S}_d$

Figure 6.1 compares graphically the power of five different tests for trivariate independence based respectively on statistics B_n, L_n, M_n, T_n and W_n. These comparisons were carried out at the 5% level for four parametric classes of three-dimensional copula alternatives considered in Section 6.3, namely

(a) The equicorrelated Gaussian copula;

(b) the Farlie–Gumbel–Morgensten (FGM) copula;

(c) the Frank or Ali–Mikhail–Haq (AMH) copula;

(d) the Clayton or Gumbel–Barnett (GB) copula.

In view of the considerations made in Sections 6.3 and 6.4, the choice of representative within each class is irrelevant for the following asymptotic local power comparisons.

The local power curves corresponding to L_n, M_n and W_n were evaluated by numerical integration of the formulas given in Propositions 6.4–6.6. In each case, the integral was computed by the trapezoidal rule on a domain of the form $[0, K]$ for suitably large K; all infinite sums were truncated to 40 terms in each index to insure numerical accuracy. As for the local power curves associated with B_n and T_n, they were obtained through Monte Carlo simulation, using 10,000 repetitions for each point along the curve.

From Figure 6.1, it can be seen that the test based on the Cramér–von Mises statistic B_n is best for the Clayton/Gumbel–Barnett models, in the neighborhood of independence. It is also close to optimal for the Frank/Ali–Mikhail–Haq models. but its behavior is much less satisfactory for the Farlie–Gumbel–Morgenstern, and especially for the equicorrelated Gaussian dependence structure.

Leaving B_n alone, the smallest local power is yielded by the dependogram statistic M_n for all structures considered, except the Farlie–Gumbel–Morgenstern. In the latter model, the linear combination statistic L_n is the least powerful locally. This may be explained by the fact that this procedure assigns equal weights to the three $B_{A,n}$ statistics of size $|A| = 2$ and to the single $B_{A,n}$ statistic of size $|A| = 3$, whereas only the latter holds any power in detecting dependence, as observed in Section 6.3.2. In contrast, note the excellent performance of W_n, which weighs the $B_{A,n}$'s proportionally to $\pi^{2|A|}$.

Given that W_n is an approximation of T_n, as seen in Section 6.4.3, it is not surprising that their behavior be similar in all models. While Fisher's procedure T_n ranks first in the trivariate normal model, there does not seem to be much to choose between T_n and W_n in the three other models considered.
Figure 6.1: Comparative local power curves of B_n (—), T_n (– –), M_n (· · ·), L_n (– · –), and W_n (– · · · –) for four classes of trivariate copula alternatives: equicorrelated Gaussian (upper left), FGM (upper right), Frank/AMH (lower left), and Clayton/GB copula (lower right).

6.5.3 Power comparisons for statistics based on sets A with $|A| = 2$

In Section 6.3, it was seen that for equicorrelated Gaussian and Clayton/Gumbel–Barnett contiguous alternatives, the drift function q_A vanishes when $|A| > 2$. As a consequence, the asymptotic distribution of $B_{A,n}$ under Q_n is then the same as that under P_n, for $|A| > 2$. A loss in efficiency may thus be expected to incur when testing for multivariate independence within these models whenever a test statistic combines all possible $B_{A,n}$, rather than only those for which $|A| = 2$.

For such types of alternatives, potentially more efficient procedures could possibly be based on

$$L_{n,2} = \sum_{|A|=2} B_{A,n}, \quad M_{n,2} = \max_{|A|=2} \left\{ \frac{B_{A,n}}{q_2(\alpha')} \right\}$$

and

$$T_{n,2} = \sum_{|A|=2} \log \left\{ 1 - F_{A,n}(B_{A,n}) \right\},$$

where $\alpha' = 1 - (1-\alpha)^{2/d(d-1)}$ and $q_2(\alpha')$ is given in Table 6.2. Analytic expressions for the asymptotic local power curves of these statistics can be derived from straightforward adaptations of Propositions 6.4–6.6. Table 6.2 gives the critical values, at the 5% level, of the tests based on these statistics, based on Monte Carlo simulations. These values were used in Figure 2 to compare the local power curves of these statistics to their analogous version based on all $|A| > 1$.

In the case $d = 3$, only the statistic $B_{A,n}$ with $|A| = 3$ has no local power in testing for independence in the equicorrelated Gaussian model. The loss of power caused by its inclusion is most apparent for the dependogram statistic M_n, as illustrated in the left panel of Figure 6.2. The right panel of the same figure shows that the loss is much less dramatic (if indeed there is any) for T_n compared to $T_{n,2}$. This was perhaps to be expected, considering that Fisher's test based on T_n is probably close to optimal in this context.

Figure 6.2: Local power of two test statistics for independence in the trivariate, equicorrelated Gaussian model; left panel: $M_{n,2}$ (solid line) versus M_n (broken line); right panel: $T_{n,2}$ (solid line) versus T_n (broken line).

6.6 Asymptotic efficiency results

Given two competing test statistics $S_{n,1}$ and $S_{n,2}$, let

$$\beta_{S_i}(\alpha, \delta) = \lim_{n \to \infty} \mathrm{P}\left\{ S_{n,i} > q_{S_i}(\alpha) \mid Q_n \right\}, \quad i = 1, 2.$$

To compare their asymptotic efficiency, one may use the ratio

$$e_{12}(\alpha) = \lim_{\delta \to 0} \frac{\beta_{S_1}(\alpha, \delta) - \alpha}{\beta_{S_2}(\alpha, \delta) - \alpha} \tag{6.15}$$

of the slopes of their respective power curves in a positive neighborhood of $\delta = 0$. As mentioned by Genest et al. (2005a), this measure generalizes that of Pitman in the case of asymptotically normally distributed statistics.

Analytic expressions can be obtained for the slope in $\delta = 0$ of the local power curve of the tests of independence based on L_n, W_n and M_n. They are given in the following proposition, whose proof depends on Proposition D.2 to be found in Appendix D. In the sequel, let β_A denote the local power function associated with the statistic $B_{A,n}$.

Proposition 6.7. *Let $h_{w,|A|}$, $h_{w,L}$ and $h_{w,W}$ be the respective densities of*

$$\mathbb{Q}_A + w\chi_2^2, \quad \sum_{|A|>1} \mathbb{Q}_A + w\chi_2^2, \quad \text{and} \quad \sum_{|A|>1} \pi^{2|A|}\mathbb{Q}_A + w\chi_2^2,$$

where in each case, the summands are taken to be mutually independent. Then, as $\delta \to 0$, one has

(i) $\delta^{-2}\{\beta_A(\alpha,\delta) - \beta_A(\alpha,0)\} \to \sum_{\gamma\in\mathbb{N}^{|A|}} \lambda_\gamma I_{\gamma,A}^2 h_{\lambda_\gamma,A}\left\{q_{|A|}(\alpha)\right\}$;

(ii) $\delta^{-2}\{\beta_L(\alpha,\delta) - \beta_L(\alpha,0)\} \to \sum_{|A|>1}\sum_{\gamma\in\mathbb{N}^{|A|}} \lambda_\gamma I_{\gamma,A}^2 h_{\lambda_\gamma,L}\left\{q_L(\alpha)\right\}$;

(iii) $\delta^{-2}\{\beta_W(\alpha,\delta) - \beta_W(\alpha,0)\} \to \sum_{|A|>1}\sum_{\gamma\in\mathbb{N}^{|A|}} \lambda_\gamma \pi^{2|A|} I_{\gamma,A}^2 h_{\lambda_\gamma,W}\left\{q_W(\alpha)\right\}$;

(iv) *given α' defined as in* (6.13),

$$\delta^{-2}\{\beta_M(\alpha,\delta)-\beta_M(\alpha,0)\} \to \left(\frac{1-\alpha}{1-\alpha'}\right)\sum_{|A|>1}\sum_{\gamma\in\mathbb{N}^{|A|}} \lambda_\gamma I_{\gamma,A}^2 h_{\lambda_\gamma,|A|}\left\{q_{|A|}(\alpha')\right\}.$$

Proof. Recalling that

$$\mathbb{B}_A = \sum_{\gamma\in\mathbb{N}^{|A|}} \lambda_\gamma \left(Z_\gamma + \delta I_{\gamma,A}\right)^2, \quad L = \sum_{|A|>1} \mathbb{B}_A \quad \text{and} \quad W = \sum_{|A|>1} \pi^{2|A|}\mathbb{B}_A,$$

results (i), (ii) and (iii) are derived easily using Proposition D.2 in Appendix D. Now using (i), one has that

$$\beta_A(\alpha,\delta) = \beta_A(\alpha,0) + \delta^2 \ell_A(\alpha) + o(\delta^2),$$

where

$$\ell_A(\alpha) = \sum_{\gamma \in \mathbb{N}^{|A|}} \lambda_\gamma I^2_{\gamma,A} h_{\lambda_\gamma,|A|} \left\{ q_{|A|}(\alpha) \right\}.$$

One can then write

$$\begin{aligned}
\beta_M(\alpha,\delta) - \beta_M(\alpha,0) &= \prod_{|A|>1} \{1-\beta_A(\alpha',0)\} - \prod_{|A|>1} \{1-\beta_A(\alpha',\delta)\} \\
&= \delta^2 \sum_{|A|>1} \ell_A(\alpha') \prod_{|D|>1, D\neq A} \{1-\beta_D(\alpha',0)\} + o(\delta^2) \\
&= \left(\frac{1-\alpha}{1-\alpha'}\right) \delta^2 \sum_{|A|>1} \ell_A(\alpha') + o(\delta^2),
\end{aligned}$$

from which (iv) follows. ♢

The above result, coupled with Equation (6.15), makes it possible to determine the local asymptotic efficiencies of statistics L_n, M_n and W_n. These values are reported in Table 6.3, along with those corresponding to statistics $L_{n,2}$ and $M_{n,2}$, for which a simple adaptation of Proposition 6.7 must be called upon.

The results in Table 6.3 are generally in line with those reported in Section 6.5, bearing in mind that for lack of analytical expressions for their local power curves, the Cramér–von Mises and Fisher test procedures, respectively based on B_n and T_n, could not be included in the efficiency comparisons.

The following noteworthy observations are offered as concluding remarks:

(a) Among trivariate tests of independence based on L_n, M_n and W_n, which involve all $B_{A,n}$'s, L_n was most efficient and M_n was least efficient in all models considered, except the Farlie–Gumbel–Morgenstern class of copulas. In the latter case, M_n was by far the best choice.

(b) Tests based on $L_{n,2}$ or $M_{n,2}$ were totally inefficient when dependence entered through a Farlie–Gumbel–Morgenstern system of trivariate copulas. This was to be expected, because in this case, only the term involving $B_{A,n}$ with $|A| = 3$ has a limiting distribution that differs under the null and under the alternative. Tests based on L_n and W_n did hardly better.

(c) For the other three models, a loss in efficiency was observed when going from $L_{n,2}$ to L_n, as well as when going from $M_{n,2}$ to M_n. In the equicorrelated Gaussian and Clayton/Gumbel–Barnett cases, this was expected, as it was argued earlier that the inclusion of statistics $B_{A,n}$ with $|A| > 2$ is then likely to dilute the power of the overall procedure. An explanation for the same occurrence in the Frank/Ali–Mikhail–Haq models is still lacking.

If nothing else, this study provides a new illustration of the truism that no single procedure could ever be declared best for testing at once against all forms of multivariate dependence. More importantly, however, the results reported herein shed new light on unsuspected communalities among classes of dependence models that may be superficially perceived as quite different. The accumulation of evidence from this and similar investigations may eventually lead to new typologies for dependence. Needless to say, much remains to be done before this goal can be achieved.

Table 6.1: Approximation of the critical values of the tests based on B_n, L_n, W_n, M_n and T_n, at the 5% level

	$d=3$	$d=4$	$d=5$
$q_B(0.05)$	0.056689	0.041238	0.025486
$q_L(0.05)$	0.140450	0.260018	0.420459
$q_W(0.05)$	18.504026	50.545070	134.387563
$q_2(\alpha')$	0.085180	0.101258	0.127307
$q_3(\alpha')$	0.010061	0.011855	0.013260
$q_4(\alpha')$	——	0.001591	0.001746
$q_5(\alpha')$	——	——	0.000218
$q_T(0.05)$	15.312305	35.090485	71.008876

Table 6.2: Approximation of the critical values of the tests based on $L_{n,2}$, $M_{n,2}$ and $T_{n,2}$, at the 5% level

	$d=3$	$d=4$	$d=5$
$q_{L_2}(0.05)$	0.135621	0.239170	0.368479
$q_2(\alpha')$	0.074794	0.090201	0.097137
$q_{T_2}(0.05)$	12.20343	20.32429	29.63573

Table 6.3: Local asymptotic relative efficiency of Cramér–von Mises type tests for trivariate independence

Alternative trivariate model	Best statistic among L_n, M_n, W_n, $L_{n,2}$, $M_{n,2}$ for the chosen model	Relative efficiency of the statistic L_n with respect to the best	$L_{n,2}$	M_n	$M_{n,2}$	W_n
Equicorrelated Gaussian	$L_{n,2}$	98.56	100.0	44.66	88.78	79.45
Farlie–Gumbel–Morgenstern	M_n	3.71	0	100.0	0	32.95
Frank Ali–Mikhail–Haq	$L_{n,2}$	99.34	100.0	65.28	88.92	86.09
Clayton Gumbel–Barnett	$L_{n,2}$	98.55	100.0	43.27	87.21	79.79

CHAPITRE 7

BOUNDS ON THE VALUE-AT-RISK FOR THE SUM OF POSSIBLY DEPENDENT RISKS

Résumé

Dans ce dernier article, des bornes inférieures et supérieures explicites sur la valeur-à-risque (VaR) de la somme de risques sont obtenues dans la situation où on ne dispose que d'une information partielle à propos de la structure de dépendance et des marges. Quand les distributions marginales sont connues, une reformulation d'un résultat de Embrechts et coll. (2003), accompagnée de quelques restrictions sur les densités, permet de calculer des bornes explicites pour la VaR sous plusieurs scénarios de dépendance. Dans le cas où seules les moyennes et les variances des risques sont fournies, des bornes explicites sont obtenues en optimisant sur toutes les valeurs possibles de la matrice de corrélation associée au vecteur des risques. Des exemples analytiques et numériques permettent d'apprécier la qualité des différentes bornes dérivées tout au long de cet article.

Abstract

In this paper, explicit lower and upper bounds on the Value-at-Risk (VaR) for the sum of possibly dependent risks are derived when only partial information is available about the dependence structure and the marginal behaviors. When the marginal distributions are known, a reformulation of a result of Embrechts et al. (2003) along with some restrictions on the densities allow to compute explicit bounds for VaR under various dependence scenarios. In the case when only the means and the variances of the risks are available, explicit bounds are obtained from an optimization over all possible values of the correlation matrix associated to the vector of risks. Analytical and numerical investigations are presented in order to appreciate the quality of these bounds.

7.1 Introduction

In finance and in actuarial science, Value-at-Risk measures, at a predetermined confidence level and under normal market conditions, the worst loss that an institution can suffer over a given time interval. In risk management, it is thus an important tool since it allows to quantify the volatility of a company's assets.

In theoretical terms, the Value-at-Risk (VaR) of a random variable at level α is simply defined as the α-quantile of its distribution. Due to its computational simplicity and for some regularity reasons, Value-at-Risk remains one of the most popular measure of risk despite the fact that it has been severely criticized for not being coherent. Since this means that VaR is not sub-additive, the risk associated to a given portfolio can be larger than the sum of the stand-alone risks of its components when measured by VaR. To this extent, see, for example, Artzner et al. (1999). In order to circumvent this problem, Embrechts et al. (2003) use the concept of copula to bound the Value-at-Risk of the sum of n risks when the marginal distributions are known.

The first purpose of this article is to provide explicit expressions for the bounds proposed by these authors in many distributional cases. This is achieved at the price of some restrictions on the densities of the risks. The second goal of the paper is to derive bounds when only the first two moments of the risks are known.

The paper is organized as follows. In Section 7.2, the necessary definitions and notations are enounced. In Section 7.3, explicit expressions for the lower and the upper bounds are obtained under some restrictions on the associated densities of the risks. In Section 7.4, the situation when only the means and the variances of the risks are known is treated. A reformulation, in a Value-at-Risk context, of the univariate extreme distributions of Kaas & Goovaerts (1986) leads to lower and upper bounds for the Value-at-Risk of the sum of n risks by optimizing over all possible values of the associated correlation matrix. This method is shown to be efficient when compared to another approach using copula theory. Finally, Section 7.5 is devoted to the comparison of the various bounds encountered in the paper.

7.2 Preliminaries

In this section, the necessary definitions and results to be used in the remaining of the paper will be stated.

Definition 7.1. *Let* $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ *be the extended real line and define* $\inf \emptyset = -\infty$*. The generalized left continuous inverse of a nondecreasing function* $f : \mathbb{R} \to \bar{\mathbb{R}}$ *is the function* $f^{-1} : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ *such that*

$$f^{-1}(t) = \inf \{s \in \mathbb{R} \mid f(s) \geq t\} .$$

From the latter definition, one has that f^{-1} is nondecreasing. In addition, if f is right continuous and $f^{-1}(t) < \infty$, then $f(s) \leq t$ implies that $f^{-1}(t) \geq s$.

Definition 7.1 is now used to define the Value-at-Risk of a single random variable.

Definition 7.2. *Let F be the (right continuous) distribution function of a random variable X. Then, the Value-at-Risk of X at level α is defined as*

$$\mathrm{VaR}_\alpha(X) = F^{-1}(\alpha).$$

In many financial applications, one often wants to evaluate the risk level of a portfolio of $n \geq 2$ possibly dependent risks. This calls for the study of the dependence structure among the risks. In the modern theory about dependence, this is accomplished via the use of copulas. For an excellent expository, see Nelsen (1999). Formally, let the multivariate distribution function of a random vector $X = (X_1, \ldots, X_n)$ be defined as $H(x_1, \ldots, x_n) = \mathrm{P}\,(X_1 \leq x_1, \ldots, X_n \leq x_n)$, and denote by $F_1, \ldots, F_n$ its associated marginal distributions. The theorem stated next, due to Sklar (1959), enables H to be linked with $F_1, \ldots, F_n$ through a distribution function $C : [0,1]^n \to [0,1]$ with uniform marginals, called a copula.

Theorem 7.1. *If H is a multivariate distribution function whose univariate marginals are $F_1, \ldots, F_n$, there exists a multidimensional copula C such that*

$$H(x_1, \ldots, x_n) = C\left\{F_1(x_1), \ldots, F_n(x_n)\right\}.$$

If $F_1, \ldots, F_n$ are continuous, then C is unique.

An interesting feature of copulas is that C contains all the information about the dependence structure of $(X_1, \ldots, X_n)$. For example, the theoretical value of dependence measures like Kendall's tau and Spearman's rho depend only on the copula underlying a given population.

The following definition gives the copula analogue of the notion of exchangeable random variables. This property is shared by many well-known members of copula families.

Definition 7.3. *A copula C is exchangeable if*

$$C(u_{\tau(1)}, \ldots, u_{\tau(n)}) = C(u_1, \ldots, u_n)$$

for any permutation τ of the set $\{1, \ldots, n\}$. When $n = 2$, C is said to be symmetric.

Now the definition of the dual associated to a given copula is provided. This notion will prove crucial throughout this work.

Definition 7.4. *Let C be the distribution function of a vector $(U_1, \ldots, U_n)$ with uniform marginals. Then, the dual of C is defined as*

$$C^d(u_1, \ldots, u_n) = \mathrm{P}\left(\bigcup_{i=1}^{n} \{U_i \leq u_i\}\right).$$

It is possible to order copulas by comparing them pointwise. Explicitly, let C_1 and C_2 be n-variate copulas such that for all $u = (u_1, \ldots, u_n) \in [0,1]^n$, the inequality $C_1(u) \leq C_2(u)$ holds. It is then said that C_1 *is smaller than* C_2, termed $C_1 \leq C_2$. A useful result is that any copula C lies between the lower and upper Fréchet–Hoeffding bounds. Specifically, it is always true that $W \leq C \leq M$, where

$$W(u) = \max\left(\sum_{i=1}^{n} u_i - n + 1,\, 0\right) \quad \text{and} \quad M(u) = \min_{1 \leq i \leq n} u_i.$$

While M is a copula in any dimensions, W fails to be a distribution function when $n > 2$. When M is the underlying copula of a vector $X = (X_1, \ldots, X_n)$, the components of X are said to be *comonotonic*. In that latter case, there is a random variable U uniformly distributed on $(0, 1)$ such that $X_i = F_i^{-1}(U)$, $1 \leq i \leq n$. For more details on comonotonicity in the actuarial science and finance, see Dhaene et al. (2000) or Dhaene et al. (2002).

Finally, define $\Pi(u) = u_1 \cdots u_n$ to be the copula associated to multivariate independence. The components of a random vector X with underlying copula C are said to be in *positive lower orthant dependence* (PLOD) if $C \geq \Pi$ and in *positive upper orthant dependence* (PUOD) when $C^d \leq \Pi^d$. If X is both PLOD and PUOD, it is said that X is positively orthant dependent (POD). In the bivariate case, these notions are equivalent, and X is said to be *positive quadrant dependent* (PQD). For a testing procedure that checks wether the components of a random vector are PQD, see Scaillet (2005).

7.3 Bounds when the marginal distributions are known

Consider the risks $X_1, \ldots, X_n$, that is n nonnegative random variables with known continuous distribution functions $F_1, \ldots, F_n$. It is assumed throughout this section that the copula C underlying the distribution of $(X_1, \ldots, X_n)$ is unknown. It is supposed, however, that partial information is available about C, namely that there is copulas C_L and C_U such that $C \geq C_L$ and

$C^d \leq C_U^d$.

Now denote by F_S the distribution function of $S = X_1 + \cdots + X_n$. In order to derive stochastic bounds on the Value-at-Risk of S, an n-variate analogue of a result due to Makarov (1981) and independently found by Rüschendorf (1982) will be recalled. The multivariate version presented herein can be found in Cossette et al. (2002). Explicitly, one has that $\underline{F}(s) \leq F_S(s) \leq \overline{F}(s)$, where

$$\underline{F}(s) = \sup_{u_1 + \cdots + u_n = s} C_L \left\{ F_1(u_1), \ldots, F_n(u_n) \right\}$$

and

$$\overline{F}(s) = \inf_{u_1 + \cdots + u_n = s} C_U^d \left\{ F_1(u_1), \ldots, F_n(u_n) \right\}.$$

Note in passing that $\underline{F}$ and $\overline{F}$ are themselves distribution functions. Frank et al. (1987) proved the best-possible nature of these bounds, while Williamson & Downs (1990) translated these results into bounds for the Value-at-Risk of the sum of two risks using the duality principle. The n-dimensional formulation of this result is stated formally in the next theorem. This is in fact a special case of Theorem 3.1 of Embrechts et al. (2003), where the Value-at-Risk of a function $\psi(x_1, \ldots, x_n)$ of n dependent risks was treated, applying the duality principle of Frank & Schweizer (1979). In all that follows, $\mathrm{VaR}_\alpha(S)$ stands for the Value-at-Risk, at level α, of the sum of n risks.

Theorem 7.2. *Let $X_1, \ldots, X_n$ be n risks with respective continuous marginal distribution functions $F_1, \ldots, F_n$. If the unknown copula C of $(X_1, \ldots, X_n)$ is such that $C \geq C_L$ and $C^d \leq C_U^d$ for some copulas C_L and C_U, then*

$$\underline{\mathrm{VaR}}_{C_U}(\alpha) \leq \mathrm{VaR}_\alpha(S) \leq \overline{\mathrm{VaR}}_{C_L}(\alpha),$$

where

$$\underline{\mathrm{VaR}}_{C_U}(\alpha) = \sup_{C_U^d(u_1,\ldots,u_n)=\alpha} \sum_{i=1}^{n} F_i^{-1}(u_i) \tag{7.1}$$

and

$$\overline{\mathrm{VaR}}_{C_L}(\alpha) = \inf_{C_L(u_1,\ldots,u_n)=\alpha} \sum_{i=1}^{n} F_i^{-1}(u_i). \tag{7.2}$$

For fixed marginal distribution functions $F_1, \ldots, F_n$, the dependence scenario leading to the worst possible Value-at-Risk is not attained under comonotonicity of the risks. In other words, there exist dependence structures such that $\mathrm{VaR}_\alpha(S)$ strictly exceeds the Value-at-Risk of n comonotonic risks, which can be seen to be equal to$F_1^{-1}(\alpha) + \cdots + F_n^{-1}(\alpha)$.

For the reason cited above, Value-at-Risk is not a coherent risk measure. The *worst-case* copula is rather given by

$$C_\alpha(u) = \begin{cases} \max\left\{C_L(u), \alpha\right\}, & u \in [\alpha, 1]^n \\ M(u), & \text{otherwise.} \end{cases}$$

This result was shown by Frank et al. (1987) and by Rüschendorf (1982) when $n = 2$ and $C_L = W$. The copula C_α have recently been investigated by Embrechts et al. (2005), where many interesting graphical interpretations are provided.

Solutions for optimization problems like those of equations (7.1) and (7.2) are commonly computed using the Lagrange multiplier. For example, the solution of (7.2) is obtained by solving the system

$$\left(F_i^{-1}\right)'(u_i) = \lambda \frac{\partial}{\partial u_i} C_L(u_1, \ldots, u_n), \quad 1 \leq i \leq n \quad \text{and} \quad C_L(u_1, \ldots, u_n) = \alpha.$$

However, an approach that will prove easier to handle in the sequel is to reformulate (7.1) and (7.2) in order to reduce the problem to an optimization with $n-1$ variables. For that purpose, let $u^{-n} = (u_1, \ldots, u_{n-1})$ be the vector obtained by removing the n-th component of $u = (u_1, \ldots, u_n)$. Then, for u^{-n} fixed, introduce the nondecreasing functions

$$x \mapsto C_{u^{-n}}(x) = C(u^{-n}, x) \quad \text{and} \quad x \mapsto C^d_{u^{-n}}(x) = C^d(u^{-n}, x)$$

and denote by $C^{-1}_{u^{-n}}$, $(C^d_{u^{-n}})^{-1}$ their respective generalized left continuous inverse as is described in Definition 7.1. The following straightforward adaptation of Theorem 7.2 can now be stated.

Proposition 7.1. *Let $X_1, \ldots, X_n$ be n risks with respective continuous marginal distribution functions $F_1, \ldots, F_n$. If the unknown copula C of the vector $(X_1, \ldots, X_n)$ is such that $C \geq C_L$ and $C^d \leq C^d_U$ for some copulas C_L and C_U, then*

$$\underline{\mathrm{VaR}}_{C_U}(\alpha) \leq \mathrm{VaR}_\alpha(S) \leq \overline{\mathrm{VaR}}_{C_L}(\alpha),$$

where

$$\underline{\mathrm{VaR}}_{C_U}(\alpha) = \sup_{C^d_U(u^{-n},0) \leq \alpha} \left[\sum_{i=1}^{n-1} F_i^{-1}(u_i) + F_n^{-1}\left\{ \left(C^d_{U,u^{-n}}\right)^{-1}(\alpha) \right\} \right] \tag{7.3}$$

and

$$\overline{\mathrm{VaR}}_{C_L}(\alpha) = \inf_{C_L(u^{-n},1) \geq \alpha} \left[\sum_{i=1}^{n-1} F_i^{-1}(u_i) + F_n^{-1}\left\{ C^{-1}_{L,u^{-n}}(\alpha) \right\} \right]. \tag{7.4}$$

In practical situations, the dependence structure of $(X_1, \ldots, X_n)$ is often unknown. However, for any copula C, the inequalities $C(u) \geq W(u)$ and

$C^d(u) \leq \widetilde{W}^d(u) = \min(1, u_1 + \cdots + u_n)$ hold. Note that $\widetilde{W}^d = W^d$ only for $n = 2$. Hence, in view of Proposition 7.1, it is always true that

$$\underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha) \leq \mathrm{VaR}_\alpha(S) \leq \overline{\mathrm{VaR}}_W(\alpha),$$

even if W is not a copula and $\widetilde{W}^d$ is not the dual of a copula when $n > 2$. In fact, only the property that C_L and C_U^d are increasing in each of their arguments was necessary to establish Theorem 7.2, and as a consequence Proposition 7.1.

These bounds can potentially be tightened whenever additional information tells that there exist copulas C_0 and C_1 such that $C \geq C_0 > W$ and $C^d \leq C_1^d < \widetilde{W}^d$. Indeed, since this implies that

$$C_{0,u^{-n}}^{-1}(\alpha) < W_{u^{-n}}^{-1}(\alpha) \quad \text{and} \quad (C_{1,u^{-n}}^d)^{-1}(\alpha) > (\widetilde{W}_{u^{-n}}^d)^{-1}(\alpha),$$

one concludes from the fact that F_n^{-1} is nondecreasing and the optimization regions are larger, that $\underline{\mathrm{VaR}}_{C_1}(\alpha) \geq \underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha)$ and $\overline{\mathrm{VaR}}_{C_0}(\alpha) \leq \overline{\mathrm{VaR}}_W(\alpha)$. In the bivariate case, the knowledge that $C \geq C_0$ can lead to a simultaneous improvement of $\underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha) = \underline{\mathrm{VaR}}_W(\alpha)$ and $\overline{\mathrm{VaR}}_W(\alpha)$ since it implies that $C^d \leq C_0^d$.

Now in order to apply the arguments in the above discussion, assume that $X_1, \ldots, X_n$ are positively lower orthant dependent (PLOD), that is $C \geq \Pi$. In this case, the possibly improved upper bound is $\overline{\mathrm{VaR}}_\Pi(\alpha)$. In other contexts, it can be supposed that the risks are in positive upper orthant dependence (PUOD), which means that $C^d \leq \Pi^d$. This can lead to a better lower bound, namely $\underline{\mathrm{VaR}}_\Pi(\alpha)$. However, these assumptions of PLOD and PUOD risks are rather imprudent in VaR-based risk management.

Remark 7.6. *Interestingly, the bounds* (7.3) *and* (7.4) *of Proposition 7.1 cannot be improved even if available information tells that* C *is bounded above. For example, no improvement is achieved even if it is known that the copula* C *of* $(X_1, \ldots, X_n)$ *satisfies* $C \leq \Pi$.

Unfortunately, explicit solutions for (7.3) or (7.4) are not always available. One then has to rely on numerical solutions in these situations. However, it will be seen that easily computable expressions can arise by making assumptions about the densities of the risks, which ensure that the function to be optimized is convex.

The next proposition generalizes previous findings made by Embrechts et al. (2002), where $\overline{\text{VaR}}_W(\alpha)$ was computed for the sum of two identically distributed Pareto and Gamma risks X_1 and X_2. The result presented herein holds true whenever the common density f of n risks is non-increasing above a certain threshold and applies to any exchangeable copula C_L such that $C \geq C_L$. Before stating it, suppose that f is differentiable.

Proposition 7.2. *Let* $X_1, \ldots X_n$ *be* n *risks with common distribution function* F *and unknown copula* C*, and suppose hypotheses A1 and A2 below hold true.*

A1. There exists $x^\star \in \mathbb{R}$ *such that* $f(x) = dF(x)/dx$ *is non-increasing for all* $x \geq x^\star$.

A2. There is an exchangeable copula C_L *such that* $C \geq C_L$ *and*

$$\frac{\partial^2}{\partial u_i \partial u_j} C^{-1}_{L,u^{-n}} \geq 0 \quad \textit{for any} \;\; 1 \leq i, j \leq n-1.$$

Then, for $\alpha \geq F(x^\star)$*, one has that* $\mathrm{VaR}_\alpha(S) \leq \overline{\mathrm{VaR}}_{C_L}(\alpha)$*, where*

$$\overline{\mathrm{VaR}}_{C_L}(\alpha) = nF^{-1}\left\{\delta_{C_L}^{-1}(\alpha)\right\}, \tag{7.5}$$

with $\delta_{C_L}(t) = C_L(t, \ldots, t)$ *being the diagonal section of* C_L*.*

Proof. Define $s_\alpha(u^{-n}) = C^{-1}_{L,u^{-n}}(\alpha)$ and $G(u) = F^{-1}(u)$. From equation (7.4) of Proposition 7.1, one can then write

$$\overline{\mathrm{VaR}}_{C_L}(\alpha) = \inf_{C_L(u^{-n},1)\geq\alpha} g(u^{-n}),$$

where

$$g(u^{-n}) = \sum_{i=1}^{n-1} G(u_i) + G \circ s_\alpha(u^{-n}).$$

By assumption A1, one deduces that $G''(u) \geq 0$ for all $u \geq F(x^\star)$, and in particular for $\alpha \leq u \leq 1$ since $\alpha \geq F(x^\star)$. Furthermore, $G'(u) = \{f \circ F^{-1}(u)\}^{-1} \geq 0$ for all u. Next, one computes that

$$\frac{\partial^2}{\partial u_i \partial u_j} g(u^{-n}) = \begin{cases} h_{ij}(u^{-n}) + G''(u_i), & i = j, \\ \\ h_{ij}(u^{-n}), & i \neq j, \end{cases}$$

where

$$h_{ij}(u^{-n}) = G'' \circ s_\alpha(u^{-n})\, s_\alpha^{(i)}(u^{-n})\, s_\alpha^{(j)}(u^{-n}) + G' \circ s_\alpha(u^{-n})\, s_\alpha^{(ij)}(u^{-n}),$$

$s_\alpha^{(i)} = \partial s_\alpha / \partial u_i$ and $s_\alpha^{(ij)} = \partial^2 s_\alpha / \partial u_i \partial u_j$. By hypothesis A2, $s_\alpha^{(ij)} \geq 0$. This, coupled with the fact that

$$s_\alpha^{(i)}(u^{-n}) = -\frac{C_L^{(i)}\left\{u^{-n}, s_\alpha(u^{-n})\right\}}{C_L^{(n)}\left\{u^{-n}, s_\alpha(u^{-n})\right\}} \leq 0,$$

where $C_L^{(i)}$ is the first-order partial derivative of C_L with respect to its ith component, allow to conclude that g is convex on $[\alpha, 1]^{n-1}$. A possible minimum is then attained where the first order derivatives of g vanish, that is at the point $\tilde{u}^{-n}$ such that

$$\frac{G'(\tilde{u}_i)}{s_\alpha^{(i)}(\tilde{u}^{-n})} + G' \circ s_\alpha(\tilde{u}^{-n}) = 0, \quad 1 \leq i \leq n-1. \tag{7.6}$$

A natural candidate is $\tilde{u}^{-n}$ that satisfies $\tilde{u}_i = s_\alpha(\tilde{u}^{-n}),\ 1 \leq i \leq n-1$. This would imply that $\tilde{u}_1 = \cdots = \tilde{u}_{n-1} = \tilde{u}$ and henceforth $\tilde{u}$ would be such that $C_L(\tilde{u}, \ldots, \tilde{u}) = \delta_{C_L}(\tilde{u}) = \alpha$, or equivalently $\tilde{u} = \delta_{C_L}^{-1}(\alpha)$. This solution satisfies the equations in (7.6) since, using the exchangeability of C_L,

$$s_\alpha^{(i)}(\tilde{u}^{-n}) = -\frac{C_L^{(i)}\left\{\tilde{u}^{-n}, s_\alpha(\tilde{u}^{-n})\right\}}{C_L^{(n)}\left\{\tilde{u}^{-n}, s_\alpha(\tilde{u}^{-n})\right\}} = -\frac{C_L^{(i)}(\tilde{u}, \ldots, \tilde{u})}{C_L^{(n)}(\tilde{u}, \ldots, \tilde{u})} = -1.$$

Finally, this solution belongs to the optimization region since

$$C_L\left(\tilde{u}^{-n}, 1\right) \geq C_L\left(\tilde{u}^{-n}, \tilde{u}\right) = \delta_{C_L}(\tilde{u}) = \alpha,$$

which completes the proof. ♢

Remark 7.7. *For a distribution function F whose associated density satisfies Hypothesis A1, one obtains easily that for all $\alpha \geq F(x^\star)$,*

$$\overline{\mathrm{VaR}}_W(\alpha) = nF^{-1}\left(\frac{\alpha + n - 1}{n}\right) \quad \textit{and} \quad \overline{\mathrm{VaR}}_\Pi(\alpha) = nF^{-1}\left(\alpha^{1/n}\right),$$

since Hypothesis A2 is met for both W and Π. Interestingly, it is clear that $\overline{\mathrm{VaR}}_W(\alpha)$ exceeds the Value-at-Risk of the sum of n comonotonic risks, namely $nF^{-1}(\alpha)$.

Hypothesis A1 is fulfilled for many important models in finance. In fact, as long as A2 holds true, the conclusion of Proposition 7.2 applies for all

$\alpha \geq 0$ when the density is non-increasing on its entire domain, as is the case with the exponential and Pareto models. Even for unimodal densities, the range of α where the result holds true can be wide enough for applications. To illustrate, let $F_{a,b}$ be the distribution function of a gamma random variable with parameters a and b. The associated density is known to be non-increasing for all $x \geq x^\star = (a-1)b$. Table 7.1 provides some values of $F_{a,b}(x^\star) = F_{a,1}(a-1)$, that is the minimum values of α for which the upper bound of Proposition 7.2 is still valid.

Table 7.1: Values of $F_{a,1}(a-1)$, where $F_{a,1}$ is the distribution function of a Gamma$(a, 1)$ distribution

a	1	1.5	2	3	4	5	10	∞
α	0.000	0.199	0.264	0.323	0.353	0.371	0.413	0.500

Here, the condition $\alpha \geq F_{a,b}(x^\star)$ that appears in Proposition 7.2 is not restrictive in practice since one is usually interested in the computation of the Value-at-Risk at large values of α. Note that in the special case of the exponential distribution, that is when $a = 1$, there is no restriction on α since the associated density is non-increasing everywhere.

Assumption A2 is satisfied as well for many copulas of interest, including W and Π. More generally, one can show that A2 is true for any Archimedean copula, that is for dependence models of the form

$$C(u_1, \ldots, u_n) = \phi^{-1}\left\{\phi(u_1) + \cdots + \phi(u_n)\right\},$$

where $\phi : (0,1] \to [0,\infty)$ is a bijective generator such that

$$(-1)^i \frac{d^i}{dt^i}\phi^{-1}(t) > 0, \quad 1 \le i \le n \quad \text{and} \quad \phi(1) = 0.$$

Many widely used multivariate families fall into this category, including the Clayton–Oakes and Frank models, as well as the extreme-value Gumbel copulas. For more details on Archimedean copulas, see Chapter 4 of the monograph by Nelsen (1999).

Now an explicit expression for the lower bound to be found in Theorem 7.2 will be given when no information about the dependence structure of n risks is available. In that case, one deduces from equation (7.1) that $\mathrm{VaR}_\alpha(S)$ is bounded below by

$$\underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha) = \sup_{u_1+\cdots+u_n=\alpha} \sum_{i=1}^{n} F_i^{-1}(u_i), \tag{7.7}$$

since it is always true that $C^d \le \widetilde{W}^d = \min(1, u_1 + \cdots + u_n)$. If the densities associated to the risks are non-increasing on a given range, then the function to be maximized will be convex. This is a key requirement in the proof of the next result, where it is assumed that for all $1 \le i \le n$, the density f_i of X_i is differentiable.

Proposition 7.3. *Suppose there exist numbers $x_i^\star$, $1 \le i \le n$ such that $f_i(x)$ is non-increasing for all $x \le x_i^\star$. Then, for $\alpha \le \min\{F_1(x_1^\star), \ldots, F_n(x_n^\star)\}$, one has that*

$$\underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha) = \max_{1\le i\le n} \left\{ F_i^{-1}(\alpha) + \sum_{1\le j\ne i\le n} F_j^{-1}(0) \right\}. \tag{7.8}$$

Proof. The proof will proceed by induction. First note that by assumption, $-f_i' \circ F_i^{-1}(u) \geq 0$ for all $u \leq F_i(x_i^\star)$, so that for all $0 \leq u \leq \alpha \leq F_i(x_i^\star)$,

$$\left(F_i^{-1}(u)\right)'' = \frac{-f_i' \circ F_i^{-1}(u)}{\left\{f_i \circ F_i^{-1}(u)\right\}^3} \geq 0, \quad 1 \leq i \leq n.$$

For $n = 2$, one deduces from equation (7.7) that

$$\begin{aligned} \underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha) &= \sup_{0 \leq u \leq \alpha} \left\{F_1^{-1}(u) + F_2^{-1}(\alpha - u)\right\} \\ &= \max\left\{F_1^{-1}(\alpha) + F_2^{-1}(0),\ F_1^{-1}(0) + F_2^{-1}(\alpha)\right\}, \end{aligned}$$

where the convexity of F_1^{-1} and F_2^{-1} was used.

Now suppose equation (7.8) is true for a given $n \geq 2$ and let $u_1, \ldots, u_{n+1}$ be any nonnegative numbers that satisfy $u_1 + \cdots + u_{n+1} = \alpha$. By the induction hypothesis,

$$\sum_{i=1}^{n} F_i^{-1}(u_i) \leq \max_{1 \leq i \leq n} \left\{ F_i^{-1}(\alpha - u_{n+1}) + \sum_{1 \leq i \neq j \leq n} F_j^{-1}(0) \right\},$$

because $u_1 + \cdots + u_n = \alpha - u_{n+1}$. Hence,

$$\sum_{i=1}^{n+1} F_i^{-1}(u_i) \leq \max_{1 \leq i \leq n} \left\{ F_i^{-1}(\alpha - u_{n+1}) + F_{n+1}^{-1}(u_{n+1}) + \sum_{1 \leq j \neq i \leq n} F_j^{-1}(0) \right\},$$

and since the result holds for $n = 2$,

$$F_i^{-1}(\alpha - u_{n+1}) + F_{n+1}^{-1}(u_{n+1}) \leq \max\left\{F_i^{-1}(\alpha) + F_{n+1}^{-1}(0),\ F_{n+1}^{-1}(\alpha) + F_i^{-1}(0)\right\}.$$

It follows that

$$\sum_{i=1}^{n+1} F_i^{-1}(u_i) \leq \max_{1 \leq i \leq n} \left\{ F_i^{-1}(\alpha) + \sum_{1 \leq j \neq i \leq n+1} F_j^{-1}(0) \right\},$$

and therefore

$$\sup_{u_1+\cdots+u_{n+1}=\alpha} \sum_{i=1}^{n+1} F_i^{-1}(u_i) = \max_{1\le i\le n} \left\{ F_i^{-1}(\alpha) + \sum_{1\le j\ne i\le n+1} F_j^{-1}(0) \right\}.$$

Consequently, (7.8) is true for $n+1$. ♢

Remark 7.8. *It is not possible to obtain an analogous version of the last proposition for any copula C_U such that $C^d \le C_U^d$, since the function to be maximized will be no longer convex. However, a result similar to that of Proposition 7.2 is possible for the lower bound when it is supposed that the common density of n risks is nondecreasing for all $x \le x^\star$. The gamma model, among others, satisfy this requirement. Under the assumption that there exists an exchangeable copula C_U such that $C^d \le C_U^d$ and*

$$\frac{\partial^2}{\partial u_i \partial u_j} (C^d_{U,u^{-n}})^{-1} \le 0 \quad \textit{for any } 1 \le i,j \le n-1,$$

the function to be maximized in equation (7.3) of Proposition 7.1 is concave. It is hence established from this fact, using arguments identical to that of the proof of Proposition 7.2, that the maximum is attained at $\tilde{u}_i = (\delta^d_{C_U})^{-1}(\alpha)$, $1 \le i \le n$, where $\delta^d_{C_U}(t) = C_U^d(t,\ldots,t)$. As a consequence, for $\alpha \le F(x^\star)$,

$$\underline{\mathrm{VaR}}_{C_U}(\alpha) = nF^{-1}\left\{\delta^d_{C_U^{-1}}(\alpha)\right\}.$$

This result is however of limited application since it holds true only for small values of α.

As an illustration of the latter remark, let F by a distribution function whose associated density $f(x)$ is nondecreasing for all $x \le x^\star$. One obtains, for all $\alpha \le F(x^\star)$, that

$$\underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha) = nF^{-1}\left(\frac{\alpha}{n}\right)$$

and

$$\underline{\mathrm{VaR}}_{\Pi}(\alpha) = nF^{-1}\left\{1 - (1-\alpha)^{1/n}\right\}.$$

7.4 Bounds when the marginal distributions are unknown

In this section, lower and upper bounds for $\mathrm{VaR}_\alpha(S)$ when only the first two moments of X_i, $1 \le i \le n$ are known are proposed. The main result will make use of the univariate extremal distributions given by Kaas & Goovaerts (1986) when only the first two moments of a random variable are known. Specifically, let X be a random variable with unknown distribution function F and known moments $\mathrm{E}(X) = \mu_X > 0$ and $\mathrm{var}(X) = \sigma_X^2 > 0$. These authors showed that $\underline{F}_{\mu_X,\sigma_X}(x) \le F(x) \le \overline{F}_{\mu_X,\sigma_X}(x)$, where

$$\underline{F}_{\mu_X,\sigma_X}(x) = \begin{cases} \dfrac{\sigma_X^2}{\sigma_X^2 + (x-\mu_X)^2}, & 0 \le x \le \mu_X, \\ \\ 1, & x > \mu_X \end{cases}$$

and

$$\overline{F}_{\mu_X,\sigma_X}(x) = \begin{cases} 0, & 0 \le x \le \mu_X, \\ \\ \dfrac{\mu_X - x}{x}, & \mu_X < x \le \dfrac{\sigma_X^2 + \mu_X^2}{\mu_X}, \\ \\ \dfrac{(x-\mu_X)^2}{(x-\mu_X)^2 + \sigma_X^2}, & x > \dfrac{\sigma_X^2 + \mu_X^2}{\mu_X}. \end{cases}$$

The following proposition translates these bounds in terms of the Value-at-Risk of a single random variable X. To achieve this, it suffices to invert the previous extremal distributions. Before stating it, define on $[0,1]$ the strictly increasing function $q(x) = \sqrt{x/(1-x)}$ and let

$$g_{a,b}(x) = \{a - bq(1-x)\}\,\mathbf{1}\left(x \geq \frac{b^2}{a^2+b^2}\right)$$

and

$$h_{a,b}(x) = a + aq^2(x)\mathbf{1}\left(x \leq \frac{b^2}{a^2+b^2}\right) + bq(x)\mathbf{1}\left(x > \frac{b^2}{a^2+b^2}\right),$$

where $\mathbf{1}(\cdot)$ stands for the indicator function of a set.

Proposition 7.4. *If X is a random variable with mean μ_X and variance σ_X^2,*

$$\underline{\mathrm{VaR}}_{\mu_X,\sigma_X}(\alpha) \leq \mathrm{VaR}_\alpha(X) \leq \overline{\mathrm{VaR}}_{\mu_X,\sigma_X}(\alpha),$$

where

$$\underline{\mathrm{VaR}}_{\mu_X,\sigma_X}(\alpha) = \overline{F}^{-1}_{\mu_X,\sigma_X}(\alpha) = g_{\mu_X,\sigma_X}(\alpha)$$

and

$$\overline{\mathrm{VaR}}_{\mu_X,\sigma_X}(\alpha) = \underline{F}^{-1}_{\mu_X,\sigma_X}(\alpha) = h_{\mu_X,\sigma_X}(\alpha).$$

Now in order to derive bounds for the Value-at-Risk of $S = X_1 + \cdots + X_n$, an approach similar to that used by Genest et al. (2002) in a stop-loss premium context will be employed. Specifically, let $X_1, \ldots, X_n$ be n risks such that $\mathrm{E}(X_i) = \mu_i > 0$ and $\mathrm{var}(X_i) = \sigma_i^2 > 0$ are known. Hence, if R stands for their associated correlation matrix, the first two moments of the single random variable S are expressed as

$$\mu = \mathrm{E}(S) = \mu_1 + \cdots + \mu_n$$

and

$$\sigma^2(R) = \mathrm{var}(S) = \sum_{i=1}^{n} \sigma_i^2 + 2\sum_{i<j} \sigma_i\sigma_j R_{ij}.$$

From Proposition 7.4, possible bounds for $\mathrm{VaR}_\alpha(S)$ are then

$$\underline{\mathrm{VaR}}_{\mu,\sigma(R)}(\alpha) \le \mathrm{VaR}_\alpha(S) \le \overline{\mathrm{VaR}}_{\mu,\sigma(R)}(\alpha).$$

Since these two extremal Value-at-Risk depend on the unknown elements of R, expressions for the lower and the upper bounds that are free of R are obtained by minimizing $\underline{\mathrm{VaR}}_{\mu,\sigma(R)}$ and maximizing $\overline{\mathrm{VaR}}_{\mu,\sigma(R)}$ with respect to $R_{ij} \in [-1,1]$, $1 \le i \ne j \le n$. These optimization problems are easily handled using the fact that $\sigma^2(R)$ and $t(R) = \sigma^2(R)/\left(\sigma^2(R)+\mu^2\right)$ are strictly increasing on each of their arguments. It follows that $\underline{\mathrm{VaR}}_{\mu,\sigma(R)}$ and $\overline{\mathrm{VaR}}_{\mu,\sigma(R)}$ are respectively decreasing and increasing functions of R_{ij}, in the strict sense, so that both solutions are achieved when $R_{ij} = 1$ for all $1 \le i,j \le n$. As a consequence, the resulting bounds depend only on $\mu = \mu_1 + \cdots + \mu_n$ and $\sigma = \sigma_1 + \cdots + \sigma_n$. These new findings are summarized in the proposition below.

Proposition 7.5. *Let $X_1,\ldots,X_n$ be n risks with respective means and variances μ_i and σ_i^2. Then, $\underline{\mathrm{VaR}}_{\mu,\sigma}(\alpha) \le \mathrm{VaR}_\alpha(S) \le \overline{\mathrm{VaR}}_{\mu,\sigma}(\alpha)$, where*

$$\underline{\mathrm{VaR}}_{\mu,\sigma}(\alpha) = g_{\mu,\sigma}(\alpha) \quad \text{and} \quad \overline{\mathrm{VaR}}_{\mu,\sigma}(\alpha) = h_{\mu,\sigma}(\alpha), \tag{7.9}$$

with $\mu = \mu_1 + \cdots + \mu_n$ and $\sigma = \sigma_1 + \cdots + \sigma_n$.

Remark 7.9. *Since the optimization consists in taking the maximum values of R_{ij} in the interval $[-1,1]$, the bounds in Proposition 7.5 can be improved* *whenever additional information tells that* $R_{ij} \leq R^\star_{ij}$ *for some pair* (i,j) *such that* $-1 \leq R^\star_{ij} < 1$*, using the fact that* $\sigma(R) \leq \sigma(R^\star)$*. One only has to replace* σ *by* $\sigma(R^\star)$ *in* (7.9) *in order to obtain better bounds. For example, if it is supposed that* $R_{ij} \leq 0$ *for all* $1 \leq i \neq j \leq n$*, then* $\sigma = \sigma(I) = \sqrt{\sigma_1^2 + \cdots + \sigma_n^2}$ *can be substituted in equation* (7.9) *in order to obtain improved lower and upper bounds.*

Remark 7.10. *An alternative method consists in bounding* $\mathrm{VaR}_\alpha(S)$ *via Proposition 3.1 by replacing* F_i^{-1}*,* $1 \leq i \leq n$*, by* $\underline{\mathrm{VaR}}_{\mu_i,\sigma_i}$ *in equation* (7.3) *and by* $\overline{\mathrm{VaR}}_{\mu_i,\sigma_i}$ *in equation* (7.4)*. If the unknown copula* C *of* $(X_1,\ldots,X_n)$ *is such that* $C \geq C_L$ *and* $C^d \leq C_U^d$ *for some copulas* C_L *and* C_U*, then*

$$\underline{\mathrm{VaR}}^\star_{C_U}(\alpha) \leq \mathrm{VaR}_\alpha(S) \leq \overline{\mathrm{VaR}}^\star_{C_L}(\alpha),$$

where

$$\underline{\mathrm{VaR}}^\star_{C_U}(\alpha) = \sup_{C_U^d(u^{-n},0)\leq\alpha} \left[\sum_{i=1}^{n-1} g_{\mu_i,\sigma_i}(u_i) + g_{\mu_n,\sigma_n}\left\{ \left(C^d_{L,u^{-n}}\right)^{-1}(\alpha) \right\} \right]$$

and

$$\overline{\mathrm{VaR}}^\star_{C_L}(\alpha) = \inf_{C_L(u^{-n},1)\geq\alpha} \left[\sum_{i=1}^{n-1} h_{\mu_i,\sigma_i}(u_i) + h_{\mu_n,\sigma_n}\left\{ C^{-1}_{L,u^{-n}}(\alpha) \right\} \right].$$

The next proposition states that the upper bound constructed from the correlation-based methodology, namely $\overline{\mathrm{VaR}}_{\mu,\sigma}$, is uniformly better than the upper bound that arises from the copula-based approach described in the remark above, at least when the coefficients of variation of two risks X_1 and X_2 are equal. An exhaustive numerical investigation suggests that this result could probably be extended to the general case. Before stating it, put

$$t_i = \frac{\sigma_i^2}{\mu_i^2 + \sigma_i^2},\ i = 1,2 \quad \text{and} \quad t = \frac{(\sigma_1+\sigma_2)^2}{(\mu_1+\mu_2)^2 + (\sigma_1+\sigma_2)^2}.$$

Proposition 7.6. *Let X_1, X_2 be two risks with μ_1, μ_2, σ_1, σ_2 known and such that $\mu_1/\sigma_1 = \mu_2/\sigma_2$. If C_L is any symmetric copula such that $C \geq C_L$, then for all $0 \leq \alpha \leq 1$,*

$$\overline{\text{VaR}}_{\mu,\sigma}(\alpha) \leq \overline{\text{VaR}}^{\star}_{C_L}(\alpha).$$

Proof. By the assumption $\mu_1/\sigma_1 = \mu_2/\sigma_2$, one has that $t_1 = t_2 = t$. Furthermore, since $C^{-1}_{L,u}$ is a non-increasing function of u, one has that $C^{-1}_{L,u}(\alpha) \geq C^{-1}_{L,1}(\alpha) = \alpha$ for any $0 \leq u \leq 1$. If $\alpha \leq t$,

$$\begin{aligned}
\overline{\text{VaR}}^{\star}_{C_L}(\alpha) &= \mu + \min\Bigg[\inf_{\alpha\leq u\leq t}\left\{\mu_1 q^2(u) + \mu_2 q^2\left(C^{-1}_{L,u}(\alpha)\right)\right\}, \\
&\qquad\qquad \inf_{t<u\leq 1}\left\{\sigma_1 q(u) + \sigma_2 q\left(C^{-1}_{L,u}(\alpha)\right)\right\}\Bigg] \\
&\geq \mu + \min\left[\mu_1 q^2(\alpha) + \mu_2 q^2(\alpha),\ \sigma_1 q(\alpha) + \sigma_2 q(\alpha)\right] \\
&= \mu + \mu_1 q^2(\alpha) + \mu_2 q^2(\alpha) \\
&= \mu + \mu q^2(\alpha) \\
&= \overline{\text{VaR}}_{\mu,\sigma}(\alpha),
\end{aligned}$$

since $\sigma_i q(\alpha) = \mu_i q(t) q(\alpha) \geq \mu_i q^2(\alpha)$, while if $t < \alpha \leq 1$, one has that

$$\begin{aligned}
\overline{\text{VaR}}^{\star}_{C_L}(\alpha) &= \mu + \inf_{\alpha\leq u\leq 1}\left\{\sigma_1 q(u) + \sigma_2 q\left(C^{-1}_{L,u}(\alpha)\right)\right\} \\
&\geq \mu + \sigma_1 q(\alpha) + \sigma_2 q(\alpha) \\
&= \mu + \sigma q(\alpha) \\
&= \overline{\text{VaR}}_{\mu,\sigma}(\alpha),
\end{aligned}$$

which completes the proof. ♢

For the lower bounds derived from the correlation-based and the copula methodologies, there is not a clear answer telling which one is preferable.

Indeed, it will be seen that when $n = 2$, the lower bound that arises from the copula-based approach is the best for $0 \leq \alpha \leq t_2$, while the one that emerges by the correlation approach is better for $t_2 < \alpha \leq 1$. This suggests an improved lower bound for the Value-at-Risk for the sum of two risks by combining $\underline{\mathrm{VaR}}_{\mu,\sigma}$ and $\underline{\mathrm{VaR}}^{\star}_{C_L}$. This is the subject of the next proposition.

Proposition 7.7. *If the first two moments μ_1, μ_2, σ_1, σ_2 of X_1, X_2 are known, and if C_U is a symmetric copula such that $C^d \leq C^d_U$, a lower bound for $\mathrm{VaR}_\alpha(S)$ that is better than $\underline{\mathrm{VaR}}_{\mu,\sigma}(\alpha)$ and $\underline{\mathrm{VaR}}^{\star}_{C_U}(\alpha)$ and that does not depend on C_U is*

$$\underline{\mathrm{VaR}}^{\star\star}_{\mu_1,\mu_2,\sigma_1,\sigma_2}(\alpha) = g_{\mu_1,\sigma_1}(\alpha) + g_{\mu_2,\sigma_2}(\alpha). \tag{7.10}$$

Proof. Assume without any loss of generality that $t_1 \leq t \leq t_2$ and note that $\mu_i - \sigma_i q(1-\alpha) \geq 0$ if and only if $\alpha \geq t_i$. It will be shown that $\underline{\mathrm{VaR}}^{\star}_{C_U}(\alpha) \geq \underline{\mathrm{VaR}}_{\mu,\sigma}(\alpha)$ for $\alpha \leq t_2$ and $\underline{\mathrm{VaR}}_{\mu,\sigma}(\alpha) \leq \underline{\mathrm{VaR}}^{\star}_{C_U}(\alpha)$ for $\alpha > t_2$.

For $\alpha \leq t_2$, it is easily established that

$$g_{\mu_1,\sigma_1}(\alpha) = g_{\mu,\sigma}(\alpha) + \{\mu_1 - \sigma_1 q(1-\alpha)\}\, \mathbf{1}(t_1 \leq \alpha \leq t_2) \geq g_{\mu,\sigma}(\alpha).$$

Hence, since $(C^d_{U,u})^{-1}$ is decreasing as a function of u, one has for all $0 \leq u \leq \alpha$ that $(C^d_{U,u})^{-1}(\alpha) \leq (C^d_{U,0})^{-1}(\alpha) = \alpha \leq t_2$, so that

$$\begin{aligned}
\underline{\mathrm{VaR}}^{\star}_{C_U}(\alpha) &= \sup_{0\leq u\leq\alpha} \left[g_{\mu_1,\sigma_1}(u) + g_{\mu_2,\sigma_2}\left\{ \left(C^d_{U,u}\right)^{-1}(\alpha) \right\} \right] \\
&= \sup_{0\leq u\leq\alpha} g_{\mu_1,\sigma_1}(u) + 0 \\
&= g_{\mu_1,\sigma_1}(\alpha) \\
&\geq g_{\mu,\sigma}(\alpha) \\
&= \underline{\mathrm{VaR}}_{\mu,\sigma}(\alpha).
\end{aligned}$$

Next, when $\alpha > t_2$, one has that $\underline{\mathrm{VaR}}_{\mu,\sigma}(\alpha) = g_{\mu_1,\sigma_1}(\alpha) + g_{\mu_2,\sigma_2}(\alpha)$, and then

$$\begin{aligned}
\underline{\mathrm{VaR}}^{\star}_{C_U}(\alpha) &= \sup_{0\le u\le\alpha}\left[g_{\mu_1,\sigma_1}(u) + g_{\mu_2,\sigma_2}\left\{\left(C^d_{U,u}\right)^{-1}(\alpha)\right\}\right] \\
&\le \sup_{0\le u\le\alpha} g_{\mu_1,\sigma_1}(u) + \sup_{0\le u\le\alpha} g_{\mu_2,\sigma_2}\left\{\left(C^d_{U,u}\right)^{-1}(\alpha)\right\} \\
&= g_{\mu_1,\sigma_1}(\alpha) + g_{\mu_2,\sigma_2}(\alpha) \\
&= \underline{\mathrm{VaR}}_{\mu,\sigma}(\alpha).
\end{aligned}$$

As a consequence, a better bound is given by

$$\begin{aligned}
\underline{\mathrm{VaR}}^{\star\star}_{\mu_1,\mu_2,\sigma_1,\sigma_2}(\alpha) &= g_{\mu_1,\sigma_1}(\alpha)\mathbf{1}(\alpha\le t_2) + \left\{g_{\mu_1,\sigma_1}(\alpha) + g_{\mu_2,\sigma_2}(\alpha)\right\}\mathbf{1}(\alpha > t_2) \\
&= g_{\mu_1,\sigma_1}(\alpha) + g_{\mu_2,\sigma_2}(\alpha).
\end{aligned}$$

$\diamondsuit$

7.5 Examples

In this section, some of the bounds established in Sections 7.3 and 7.4 will be computed in the special cases when the risks are distributed as exponential and Pareto random variables. Since the densities associated to these laws are non-increasing everywhere on their domain, the results of Propositions 7.2 and 7.3 will be valid for all $0 \le \alpha \le 1$. The computations herein can be seen as extending Examples 1-2 in Denuit et al. (1999) in a Value-at-Risk context.

7.5.1 Exponential risks

Suppose $X_1, \ldots, X_n$ are distributed as shifted exponential variables, so that

$$F_i^{-1}(u) = \xi_i - \theta_i \log(1-u), \quad \theta_i > 0, \quad \xi_i \geq 0.$$

When no information is available about the dependence structure of $(X_1, \ldots, X_n)$, one deduces from Theorem 7.2 that

$$\text{VaR}_\alpha(S) \leq \overline{\text{VaR}}_W(\alpha) = \sum_{i=1}^n \xi_i + \inf_{u_1+\cdots+u_n=\alpha+n-1} \sum_{i=1}^n -\theta_i \log(1-u_i).$$

Since the function to be minimized is convex, the problem can be solved using the Lagrange multiplier method, which gives

$$\frac{\theta_i}{1-u_i} = \lambda, \quad 1 \leq i \leq n, \quad \text{and} \quad \sum_{i=1}^n u_i = \alpha + n - 1.$$

The solution to this system of equations is $u_i = 1-(1-\alpha)\theta_i \left(\theta_1 + \cdots + \theta_n\right)^{-1}$, so that

$$\overline{\text{VaR}}_W(\alpha) = \sum_{i=1}^n \left[\xi_i - \theta_i \log\left\{(1-\alpha)\frac{\theta_i}{\theta_1 + \cdots + \theta_n}\right\}\right].$$

When $F_1 = \cdots = F_n = F$, that is for $\xi_1 = \cdots = \xi_n = \xi$ and $\theta_1 = \cdots = \theta_n = \theta$, the formula above reduces to

$$\overline{\text{VaR}}_W(\alpha) = n\xi - n\theta \log\left(1 - \frac{\alpha+n-1}{n}\right) = nF^{-1}\left(\frac{\alpha+n-1}{n}\right),$$

as can be deduce from Proposition 7.2.

This upper bound can potentially be improved when it is known that $X_1, \ldots, X_n$ are in positive lower orthant dependence (PLOD). While no simple solution seems possible for $n > 2$, one has from equation (7.4) of Proposition 7.1 that an upper bound for the Value-at-Risk of the sum of two exponential risks is

$$\overline{\text{VaR}}_{\Pi}(\alpha) = \xi_1 + \xi_2 + \inf_{\alpha \le u \le 1} s(u), \quad \text{where } s(u) = -\theta_1 \log(1-u) - \theta_2 \log\left(1 - \frac{\alpha}{u}\right).$$

From the fact that $\theta_1 > 0$, $\theta_2 > 0$, $\alpha \ge 0$ and $u \ge \alpha/2$, it follows that s is a convex function since

$$s''(u) = \frac{\theta_1}{(1-u)^2} + \frac{\theta_2 \alpha}{u^2(u-\alpha)^2}(2u - \alpha) > 0.$$

Therefore, a possible minimum for s is attained for $u^\star$ such that $s'(u^\star) = 0$. A straightforward computation gives the unique solution

$$u^\star = \frac{\alpha(\theta_1 - \theta_2) + \sqrt{\alpha^2(\theta_1 - \theta_2)^2 + 4\alpha\theta_1\theta_2}}{2\theta_1}.$$

For equal distributions, it is easily seen that $u^\star = \sqrt{\alpha} = \delta_{\Pi}^{-1}(\alpha)$, in accordance with Proposition 7.2.

Finally, one deduces from Proposition 7.3 that

$$\text{VaR}_\alpha(S) \ge \underline{\text{VaR}}_{\widetilde{W}}(\alpha) = \sum_{i=1}^{n} \xi_i - \log(1-\alpha) \max_{1 \le i \le n} \theta_i.$$

Unfortunately, no simple solution seems available when additional information tells, for instance, that $C^d \le \Pi^d$.

Now if only the first two moments of $X_1, \ldots, X_n$ are known, lower and upper bounds derived from Proposition 7.5 are given by

$$\underline{\text{VaR}}_{\xi,\theta}(\alpha) = g_{\theta+\xi,\theta}(\alpha) \quad \text{and} \quad \overline{\text{VaR}}_{\xi,\theta}(\alpha) = h_{\theta+\xi,\theta}(\alpha),$$

where $\theta = \theta_1 + \cdots + \theta_n$ and $\xi = \xi_1 + \cdots + \xi_n$. In the special case of two exponential risks, Proposition 7.7 gives the improved lower bound $\underline{\text{VaR}}^{\star\star}_{\xi_1,\xi_2,\theta_1,\theta_2}(\alpha) = g_{\theta_1+\xi_1,\theta_1}(\alpha) + g_{\theta_2+\xi_2,\theta_2}(\alpha)$.

In Figure 7.1, the curves defined by $\underline{\mathrm{VaR}}^{\star\star}_{\xi_1,\xi_2,\theta_1,\theta_2}$, $\underline{\mathrm{VaR}}_{\widetilde{W}}$, $\overline{\mathrm{VaR}}_W$ and $\overline{\mathrm{VaR}}_{\xi,\lambda}$ are displayed for the case of two exponential risks with parameter values $\xi_1 = \xi_2 = 0$ and $\theta_1 = \theta_2 = 1$. These bounds are compared to the Value-at-Risk of $S = X_1 + X_2$ when the risks are supposed to be comonotonic. Table 7.2 reports some numerical values for popular levels of α. As expected, the bounds when the marginal distributions are known are much closer to the exact Value-at-Risk compared to the bounds when only the first two moments are known. Moreover, for large values of α, the upper bound $\overline{\mathrm{VaR}}_W(\alpha)$ gives a rather good approximation to $\mathrm{VaR}_\alpha(S)$ compared to the performance of $\underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha)$. In fact, one has that

$$\lim_{\alpha\to 1}\frac{\overline{\mathrm{VaR}}_W(\alpha)}{\mathrm{VaR}_\alpha(S)} = 1 \quad \text{and} \quad \lim_{\alpha\to 1}\frac{\underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha)}{\mathrm{VaR}_\alpha(S)} = 1/2.$$

Table 7.2: Numerical values, for selected levels, of the bounds on the Value-at-Risk of the sum of two exponential risks

α	$\underline{\mathrm{VaR}}^{\star\star}_{\xi_1,\xi_2,\theta_1,\theta_2}(\alpha)$	$\underline{\mathrm{VaR}}_W(\alpha)$	$\mathrm{VaR}_\alpha(S)$	$\overline{\mathrm{VaR}}_W(\alpha)$	$\overline{\mathrm{VaR}}_{\mu,\sigma}(\alpha)$
0.900	1.33	2.30	4.61	5.99	8.00
0.950	1.54	3.00	5.99	7.38	10.72
0.975	1.68	3.69	7.38	8.76	14.49
0.990	1.80	4.61	9.21	10.60	21.90
0.995	1.86	5.30	10.60	11.98	30.21

Figure 7.1: Bounds on the Value-at-Risk of the sum of two exponential risks when only the first two moments are known (broken lines) and when the marginal distributions are known (dots), compared to the exact Value-at-Risk of the sum of two comonotonic exponential risks (solid lines)

7.5.2 Pareto risks

Consider two risks X_1 and X_2 distributed as Pareto random variables with parameters $\gamma_i > 0$ and $\beta_i > 2$. In that case,

$$F_i^{-1}(u) = \gamma_i \left\{ (1-u)^{-1/\beta_i} - 1 \right\}.$$

From equation (7.4) of Proposition 7.1, one observes that

$$\overline{\mathrm{VaR}}_W(\alpha) = \inf_{\alpha \le u \le 1} h(u),$$

where

$$h(u) = \gamma_1 \left\{ (1-u)^{-1/\beta_1} - 1 \right\} + \gamma_2 \left\{ (u-\alpha)^{-1/\beta_2} - 1 \right\}.$$

Since $h''(u) \ge 0$ for all $u \in [\alpha, 1]$, a possible minimum value is the real number $u^\star$ which solves $h'(u^\star) = 0$, or equivalently

$$\frac{\gamma_1}{\beta_1}(1-u^\star)^{-1-1/\beta_1} = \frac{\gamma_2}{\beta_2}(u^\star - \alpha)^{-1-1/\beta_2}.$$

A numerical routine is needed to solve this problem in general. However, an explicit solution arises when $\beta_1 = \beta_2 = \beta$. In that case, $u^\star = (\alpha\kappa+1)/(\kappa+1)$, where $\kappa = (\gamma_1/\gamma_2)^{\beta/(\beta+1)}$, so that

$$\overline{\mathrm{VaR}}_W(\alpha) = \left(\frac{\kappa+1}{1-\alpha} \right)^{1/\beta} \left(\frac{\gamma_1}{\kappa^{1/\beta}} + \gamma_2 \right) - \gamma_1 - \gamma_2.$$

An improved upper bound when (X_1, X_2) are known to be PQD can be computed from equation (7.4). If $\beta_1 = \beta_2 = \beta$, it can be shown that

$$\overline{\mathrm{VaR}}_{\Pi}(\alpha) = \left(\gamma_1 + \kappa^{1/\beta}\gamma_2\right)\left\{(1-u^\star)^{-1/\beta} - 1\right\},$$

where

$$u^\star = \frac{1-\kappa+\sqrt{(1-\kappa)^2+4\alpha\kappa}}{2}.$$

For equal distributions, in which case $\kappa = 1$, the above formulas reduce to

$$\overline{\mathrm{VaR}}_{W}(\alpha) \;=\; 2\gamma\left\{\left(1-\frac{\alpha+1}{2}\right)^{-1/\beta} - 1\right\} = 2F^{-1}\left(\frac{\alpha+1}{2}\right)$$

and

$$\overline{\mathrm{VaR}}_{\Pi}(\alpha) \;=\; 2\gamma\left\{\left(1-\sqrt{\alpha}\right)^{-1/\beta} - 1\right\} = 2F^{-1}\left(\sqrt{\alpha}\right),$$

in accordance with the conclusion of Proposition 7.2.

For the lower bound, one has from (7.8) that

$$\underline{\mathrm{VaR}}_{\widetilde{W}}(\alpha) = \max\left[\gamma_1\left\{(1-\alpha)^{-1/\beta_1} - 1\right\}, \gamma_2\left\{(1-\alpha)^{-1/\beta_2} - 1\right\}\right].$$

Finally, lower and upper bounds for the Value-at-Risk of $S = X_1 + X_2$ are obtained from Propositions 7.5 and 7.7 when the only available information is about the first two moments of X_i, namely

$$\mu_i = \frac{\gamma_i}{\beta_i - 1} \quad \text{and} \quad \sigma_i^2 = \frac{\gamma_i^2\beta_i}{(\beta_i-1)^2(\beta_i-2)}, \quad i = 1, 2.$$

CONCLUSION

Dans ce travail, des tests d'adéquation pour des modèles de copules ont d'abord été développés. Les statistiques de tests proposées sont des fonctionnelles de type Cramér–von Mises et Kolmogorov–Smirnov calculées sur des processus empiriques dont la convergence a été obtenue sous l'hypothèse nulle d'une appartenance à une famille donnée. Les hypothèses nécessaires à l'établissement de ce résultat asymptotique ont été vérifiées pour un grand nombre de familles de copules. La technique du *bootstrap paramétrique* a été employée pour calculer les seuils asymptotiquement exacts de ces statistiques. Une étude de puissance a permis de constater que les tests basés sur une version modifiée du processus de Kendall sont généralement efficaces sous divers scénarios de dépendance. Au passage, l'expression erronée fournie par Shih (1998) pour la variance asymptotique d'un test d'adéquation a été corrigée. La possibilité de tester l'adéquation à une famille de modèles multivariés en employant une modification judicieuse du processus de copule empirique a également été abordée.

Dans un deuxième temps, la performance dans un voisinage de l'indépendance de plusieurs tests d'indépendance multivariée a été étudiée. Un résultat intéressant sur le comportement sous des hypothèses contiguës à l'indépendance du processus de copule empirique a permis d'étudier l'efficacité locale de quelques procédures de test proposées dans la littérature. Des expressions analytiques pour les courbes de puissance ont été obtenues pour plusieurs de ces tests, et une mesure qui généralise la notion d'efficacité relative asymptotique de Pitman a été proposée. Ceci a permis, en particulier, de comparer l'efficacité locale d'une statistique pour l'indépendance bivariée, inspirée de Blum et coll. (1961) et proposée par Deheuvels, à la statistique linéaire de rangs localement la plus puissante pour une contre-hypothèse choisie. De plus, l'efficacité locale de plusieurs tests d'indépendance multivariée a été étudiée. Il a été constaté que les tests obtenus d'une décomposition de Möbius du processus de copule empirique, dont quelques-uns sont étudiés par Genest & Rémillard (2004), se comportent généralement mieux que la version multivariée de la statistique de Deheuvels.

Enfin, des bornes explicites sur la valeur-à-risque (VaR) de la somme de risques dépendants ont été obtenues lorsque l'information disponible est partielle. En particulier, des expressions explicites pour les bornes étudiées par Embrechts et coll. (2003) quand la structure de dépendance relative aux risques est inconnue sont obtenues pour des risques dont les densités respectives sont monotones sur un certain intervalle. Un des résultats obtenus formalise et généralise une observation faite par Embrechts et coll. (2002) dans le cas particulier de deux risques comonotones distribués selon des lois gamma. Aussi, en utilisant une approche basée sur les corrélations similaire à celle employée par Genest et coll. (2002), des bornes sont obtenues quand seuls les deux premiers moments des risques sont connus.

Dans le futur, des tests d'adéquation basés sur des caractéristiques suffisamment discriminantes relatives à une famille de copules pourraient être développés. En particulier, pour des observations tirées de la copule normale d-variée, les variables aléatoires $T_{ni} = Z_{ni}^\top R_n^{-1} Z_{ni}$, $1 \leq i \leq n$, où

$$Z_{ni} = \big(N^{-1} \circ F_{n1}(X_{i1}), \ldots, N^{-1} \circ F_{n1}(X_{id})\big)$$

et R_n est un estimateur de la matrice de corrélation, devraient se comporter asymptotiquement comme des variables de loi khi-deux à d degrés de liberté. L'obtention de la convergence faible du processus empirique basé sur $T_{n1}, \ldots, T_{nn}$ qui, dans la terminologie de Ghoudi & Rémillard (1998, 2004), sont des pseudo-observations, permettrait de proposer des tests d'adéquation spécifiques à la structure de dépendance gaussienne. Également, pour poursuivre l'étude d'efficacité asymptotique locale, le problème de tester l'indépendance entre des vecteurs aléatoires pourrait être considéré. Finalement, des bornes explicites sur la valeur-à-risque (VaR) pour le cas plus général d'une fonction ψ de risques dépendants pourraient éventuellement être calculées, sous des hypothèses semblables à celles énoncées pour le cas de la somme.

ANNEXE A

A.1 A convergence result

This section offers a proof of the fact that under hypotheses III and IV given in Section 2.2, the process $B_n(t) = \sqrt{n}\,\{K(\theta_n, t) - K(\theta, t)\}$ is such that

$$\sup_{t\in[0,1]} \left| B_n(t) - \dot{K}(\theta,t)^\top \Theta_n \right| \xrightarrow{P} 0.$$

To see this, let $\lambda > 0$ be arbitrary. By hypothesis III, the sequence (Θ_n) is tight since it converges in law to Θ. Hence for any given $\delta > 0$, there exist $M = M_\delta \in \mathbb{R}^+$ and N_0 such that $\mathrm{P}\{\|\Theta_n\| > M\} < \delta$ for all $n \geq N_0$. For any such n, one has

$$\mathrm{P}\left\{ \sup_{t\in[0,1]} \left| B_n(t) - \dot{K}(\theta,t)^\top \Theta_n \right| > \lambda \right\}$$

$$\begin{aligned}
&\leq \; \mathrm{P}\left\{ \sup_{t\in[0,1]} \left| B_n(t) - \dot{K}(\theta,t)^\top \Theta_n \right| > \lambda, \|\Theta_n\| \leq M \right\} + \mathrm{P}\{\|\Theta_n\| > M\} \\
&< \; \mathrm{P}\left\{ \sup_{t\in[0,1]} \left| B_n(t) - \dot{K}(\theta,t)^\top \Theta_n \right| > \lambda, \|\Theta_n\| \leq M \right\} + \delta.
\end{aligned}$$

Next, the mean-value theorem implies that for any realization of Θ_n, there exists $\theta_n^\star$ with $|\theta_n^\star - \theta| \leq |\Theta_n|/\sqrt{n}$ such that $B_n(t) = \dot{K}(\theta_n^\star, t)^\top \Theta_n$. Hence, using hypothesis IV, one obtains

$$\begin{aligned}
&\lim_{n\to\infty} P\left\{\sup_{t\in[0,1]}\left|B_n(t)-\dot{K}(\theta,t)^\top\Theta_n\right|>\lambda, \|\Theta_n\|\le M\right\}\\
&\quad\le \lim_{n\to\infty} P\left\{\|\Theta_n\|\sup_{t\in[0,1]}\left\|\dot{K}(\theta_n^\star,t)-\dot{K}(\theta,t)\right\|>\lambda, \|\Theta_n\|\le M\right\}\\
&\quad\le \lim_{n\to\infty} P\left\{\sup_{\|\theta^\star-\theta\|\le M/\sqrt{n}}\ \sup_{t\in[0,1]}\left\|\dot{K}(\theta^\star,t)-\dot{K}(\theta,t)\right\|>\frac{\lambda}{M}\right\}=0.
\end{aligned}$$

Since δ can be chosen arbitrarily small, the result follows. ♢

A.2 Verification of hypothesis IV for various copula models

A.2.1 Ali–Mikhail–Haq copulas

In view of relation (2.12), $f_i(\theta,t)$ is a polynomial in t with coefficients that are non-negative whenever $\theta\in(0,1)$, so that

$$\phi_\theta^{-1}(x)=\frac{1-\theta}{e^{x(1-\theta)}-\theta}\,,\quad x>0$$

satisfies condition (2.10) for every integer $d\ge 1$, and hence is completely monotone. It is easy to see that

$$K(\theta,t)=t+t\sum_{i=1}^{d-1}\frac{p_i(\theta,t)}{(1-\theta)^i}\ \left\{\log\left(\frac{1-\theta}{t}+\theta\right)\right\}^i,$$

where $p_i(\theta,t)=f_i(\theta,t)/t$ is also a polynomial in both θ and t. Then

$$\begin{aligned}
\dot{K}(\theta,t) &= t\sum_{i=1}^{d-1}\frac{(1-\theta)\dot{p}_i(\theta,t)+ip_i(\theta,t)}{(1-\theta)^{i+1}}\ \left\{\log\left(\frac{1-\theta}{t}+\theta\right)\right\}^i\\
&\quad -\frac{t(1-t)}{1-\theta+\theta t}\sum_{i=1}^{d-1}\frac{ip_i(\theta,t)}{(1-\theta)^i}\ \left\{\log\left(\frac{1-\theta}{t}+\theta\right)\right\}^{i-1}
\end{aligned}$$

is continuous on $(-1,1)\times[0,1]$. Hence (2.6) holds for all $\theta\in(-1,1)$.

Note, however, that hypothesis IV does not generally hold at $\theta=1$. In the case $d=2$, for example,

$$0=\lim_{\theta\to1}\lim_{t\to0}\dot{K}(\theta,t)\neq\lim_{t\to0}\lim_{\theta\to1}\dot{K}(\theta,t)=1/2.$$

A.2.2 Clayton copulas

First, it is easily seen that

$$f_{i,\theta}(t)=\frac{d^i}{ds^i}\,\phi_\theta^{-1}(s)\bigg|_{s=\phi_\theta(t)}=(-1)^i q(\theta,i,1)t^{1+i\theta},$$

where $q(\theta,i,m)=\prod_{j=0}^{i-1}(m+j\theta)$ is a polynomial of degree $i-1$ in θ. It then follows that

$$K(\theta,t)=t+t\sum_{i=1}^{d-1}\left(\frac{1-t^\theta}{\theta}\right)^i\frac{q(\theta,i,1)}{i!}$$

and

$$k(\theta,t)=\left(\frac{1-t^\theta}{\theta}\right)^{d-1}\frac{q(\theta,d,1)}{(d-1)!}\,.$$

Now since

$$\int_0^1 t\left(1-t^\theta\right)^i\,\mathrm{d}t=\frac{\Gamma(2/\theta)i!}{\theta\Gamma(2/\theta+i+1)}=\frac{\theta^{i+1}i!}{\theta\prod_{j=0}^{i}(2+j\theta)}=\frac{\theta^i i!}{q(\theta,i+1,2)}\,,$$

one has, according to formula (2.7),

$$\begin{aligned}\tau(\theta)&=1-\left(\frac{2^d}{2^{d-1}-1}\right)\sum_{i=1}^{d-1}\frac{q(\theta,i,1)}{q(\theta,i+1,2)}\\&=\left(\frac{2^d}{2^{d-1}-1}\right)\frac{q(\theta,d,1)}{q(\theta,d,2)}-\frac{1}{2^{d-1}-1}\,.\end{aligned}$$

Writing

$$\log q(\theta,i,m) = \sum_{j=0}^{i-1} \log(m+j\theta),$$

it follows that

$$q'(\theta,i,m) = q(\theta,i,m) \sum_{j=1}^{i-1} \left(\frac{j}{m+j\theta} \right)$$

and

$$\dot{k}(\theta,t) = k(\theta,t) \left(\sum_{j=1}^{d-1} \frac{j}{1+j\theta} \right) - \left(\frac{1-t^\theta}{\theta} \right)^{d-2} \frac{q(\theta,d,1)}{(d-2)!} \left(\frac{t^\theta}{\theta} \log t + \frac{1-t^\theta}{\theta^2} \right),$$

which is clearly continuous for $(0,\infty)\times[0,1]$.

Note that in this case, hypothesis IV also holds true at the boundary value $\theta = 0$. For,

$$\lim_{\theta\to 0^+} \dot{K}(\theta,t) = \frac{-t(-\log t)^d}{2(d-2)!},$$

so that

$$\dot{K}(\varepsilon,t) + \frac{t(-\log t)^d}{2(d-2)!} = \frac{F(\varepsilon,t)}{\varepsilon^d},$$

where

$$\begin{aligned} F(\varepsilon,t) &= t \sum_{i=1}^{d-1} \frac{\varepsilon^{d-i-1}(1-t^\varepsilon)^{i-1}}{(i-1)!} \left\{ t^\varepsilon - \varepsilon t^\varepsilon \log t - 1 + \frac{\varepsilon(1-t^\varepsilon)}{i} \sum_{j=1}^{i-1} \frac{j}{1+j\varepsilon} \right\} \\ &\quad + \frac{t(-\varepsilon \log t)^d}{2(d-2)!}. \end{aligned}$$

Next, in view of the general fact that

$$\frac{d^p}{d\varepsilon^p} f(\varepsilon)g(\varepsilon)h(\varepsilon) = \sum_{k=0}^{p} \sum_{\ell=0}^{k} \binom{p}{k} \binom{k}{\ell} f^{(\ell)}(\varepsilon) g^{(k-\ell)}(\varepsilon) h^{(p+1-k)}(\varepsilon),$$

and since

$$\frac{\partial^d}{\partial \varepsilon^d} (1-t^\varepsilon)^i = (\log t)^d \sum_{k=1}^{i} \binom{i}{k} (-t^\varepsilon)^k k^d$$

is continuous and bounded for all $(\varepsilon, t) \in [0,1]^2$, one finds that

$$F^{(d+1)}(\varepsilon, t) = \frac{\partial^{d+1}}{\partial \varepsilon^{d+1}} F(\varepsilon, t)$$

is bounded by a constant $M > 0$ on the unit square. Hence

$$\sup_{(\varepsilon,t)\in[0,1]^2} |F(\varepsilon, t)| = \left| \int_0^{\varepsilon} \int_0^{u_{d+1}} \cdots \int_0^{u_2} F^{(d+1)}(u_1, t)\, \mathrm{d}u_1 \cdots \mathrm{d}u_{d+1} \right| \leq \frac{\varepsilon^{d+1} M}{(d+1)!}.$$

Therefore

$$\lim_{\varepsilon \to 0} \sup_{t \in [0,1]} \left| \frac{F(\varepsilon, t)}{\varepsilon^d} \right| \leq \lim_{\varepsilon \to 0} \frac{\varepsilon M}{(d+1)!} = 0.$$

Consequently, hypothesis IV is satisfied for all $\theta \in (0, \infty)$, as well as at the boundary value $\theta = 0$.

A.2.3 Frank copulas

In the light of the remark made in Section 2.2, and from the fact that

$$\left| \ddot{K}(\theta, t) \right| \leq \int_0^t \left| \ddot{k}(\theta, s) \right| \mathrm{d}s \leq \int_0^1 \left| \ddot{k}(\theta, s) \right| \mathrm{d}s,$$

it suffices to show that $\ddot{k}(\theta, s)$ is uniformly bounded for all $\theta \in \mathcal{O}$ and $s \in [0,1]$ in order to verify hypothesis IV. For that purpose, introduce the continuous and bounded function

$$\psi(x) = \log\left(\frac{1 - e^{-x}}{x} \right), \quad x > 0$$

so that

$$\phi_\theta(t) = \psi(\theta) - \psi(\theta t) - \log t.$$

Next, let $p_0(x) = x - 1$ and $p_i(x) = x(1-x)p'_{i-1}(x)$ for arbitrary integer $i \geq 1$. With this notation, the formula already derived by Barbe et al. (1996) can be written as

$$k(\theta, s) = e^{\theta t} p'_{d-2}(e^{\theta t}) \{\log t + \psi(\theta t) - \psi(\theta)\}^{d-1}/(d-1)!, \quad \theta \geq 0.$$

One can show that

$$\begin{aligned}
&\ddot{k}(\theta, s) \\
&= s\dot{k}(\theta, s) + \frac{s^2 e^{2\theta s}}{(d-1)!} \{\log s + \psi(\theta s) - \psi(\theta)\}^{d-1} \left\{2p''_{d-2}(e^{\theta s}) + e^{\theta s} p'''_{d-2}(e^{\theta s})\right\} \\
&\quad + \frac{s e^{\theta s}}{(d-2)!} \{\log s + \psi(\theta s) - \psi(\theta)\}^{d-2} \{s\psi'(\theta s) - \psi'(\theta)\} \{2e^{\theta s} p''_{d-2}(e^{\theta s}) + p'_{d-2}(e^{\theta s})\} \\
&\quad + \frac{e^{\theta s} p'_{d-2}(e^{\theta s})}{(d-3)!} \{\log s + \psi(\theta s) - \psi(\theta)\}^{d-3} \{s\psi'(\theta s) - \psi'(\theta)\}^2 \\
&\quad + \frac{e^{\theta s} p'_{d-2}(e^{\theta s})}{(d-2)!} \{\log s + \psi(\theta s) - \psi(\theta)\}^{d-2} \{s^2 \psi''(\theta s) - \psi''(\theta)\}
\end{aligned}$$

is bounded for all $(\theta, s) \in (-\infty, \infty) \times [0, 1]$ since both ψ' and ψ'' are bounded and

$$\begin{aligned}
\dot{k}(\theta, s) &= sk(\theta, s) + \frac{s e^{2\theta s} p''_{d-2}(e^{\theta s})}{(d-1)!} \{\log s + \psi(\theta s) - \psi(\theta)\}^{d-1} \\
&\quad + \frac{e^{\theta s} p'_{d-2}(e^{\theta s})}{(d-2)!} \{\log s + \psi(\theta s) - \psi(\theta)\}^{d-2} \{s\psi'(\theta s) - \psi'(\theta)\}.
\end{aligned}$$

A.2.4 Gumbel–Hougaard copulas

This copula also belongs to the family of extreme value copulas, further discussed in Section 2.3.2. From Barbe et al. (1996),

$$k(\theta, t) = \frac{p_{d-1}(-\log t)}{(d-1)!},$$

where $p_0(x) \equiv 1$ and for integer $i \geq 1$,

$$p_i(x) = (1-\theta)x \left\{p_{i-1}(x) - p'_{i-1}(x)\right\} + (\theta + i - 1) p_{i-1}(x)$$

is a polynomial of degree i in x and in θ. This corrects a typographical error on p. 207 of Barbe et al. (1996), where one should have read

$$p_i(x) = \theta x \left\{ p_{i-1}(x) - p'_{i-1}(x) \right\} + (i-\theta) p_{i-1}(x)$$

in the parameterization used there.

Writing $p_{d-1}(x) = \sum_{k=0}^{d-1} r_k(\theta) x^k$, one may thus conclude that

$$\begin{aligned} K(\theta,t) &= \int_0^t k(\theta,s)\,\mathrm{d}s \\ &= \sum_{k=0}^{d-1} \frac{r_k(\theta)}{(d-1)!} \int_0^t (-\log s)^k \,\mathrm{d}s \\ &= \frac{t}{(d-1)!} \sum_{k=0}^{d-1} k!\; r_k(\theta) \sum_{i=0}^{k} \frac{(-\log t)^i}{i!}\,, \end{aligned}$$

so that

$$\dot{K}(\theta,t) = \frac{t}{(d-1)!} \sum_{k=0}^{d-1} k!\; r'_k(\theta) \sum_{i=0}^{k} \frac{(-\log t)^i}{i!}$$

is clearly continuous on $[0,1]^2$.

A.2.5 Fréchet copulas

Note that for copulas in this class, one has

$$\begin{aligned} \dot{K}(\theta,t) &= t\frac{\dot{\zeta}(\theta,t)}{\zeta(\theta,t)} \\ &= -2\frac{t}{I(\theta,t)} + 4\frac{t^2}{I(\theta,t)\{I(\theta,t)+\theta\}} \\ &= -2\frac{t}{I(\theta,t)} + \frac{t\zeta(\theta,t)\{I(\theta,t)+\theta\}}{I(\theta,t)}\,. \end{aligned}$$

It is easy to check that $t/I(\theta,t)$ is continuous on $[0,1]^2$, whence hypothesis IV holds true on $\mathcal{O} = (0,1)$. To see that the latter is also verified at $\theta = 1$, note that $\dot{K}(\theta, t) \to -t$ as $\theta \to 1$. Putting $\delta = 1 - \varepsilon$, one gets $\dot{K}(\varepsilon, t) + t = F(\delta, t)/\delta$, where

$$F(\delta, t) = t\left\{\delta + 1 - \frac{\delta + 1}{I(1-\delta, t)}\right\}.$$

It can easily be shown that

$$\begin{aligned} &\sup_{(\delta,t)\in[0,1)\times[0,1]} \left|\ddot{K}(\delta, t)\right| \\ &\qquad = \sup_{(\delta,t)\in[0,1)\times[0,1]} \left|\frac{\{4t(1-3t) + 4\delta t(\delta + t - 3) - 4(\delta - 1)^2\}\, t}{(I_{1-\delta,t})^{5/2}}\right| \end{aligned}$$

is bounded above by some constant M. Thus,

$$\sup_{(\delta,t)\in[0,1)\times[0,1]} |F(\delta, t)| = \sup_{(\delta,t)\in[0,1)\times[0,1]} \left|\int_0^\delta \int_0^v F(u, t)\, \mathrm{d}u\, \mathrm{d}v\right| \le \frac{\delta^2 M}{2},$$

which implies that

$$\sup_{(\delta,t)\in[0,1)\times[0,1]} |F(\delta, t)/\delta| \le \delta M/2 \to 0 \quad \text{as } \varepsilon = 1 - \delta \to 1.$$

A.2.6 Farlie–Gumbel–Morgenstern copulas

Here, one has

$$\begin{aligned} \dot{K}(\theta, t) &= \int_0^t \int_s^1 \dot{h}(\theta, x, s)\, \mathrm{d}x\, \mathrm{d}s \\ &= -\int_0^t \int_s^1 \frac{\dot{r}(\theta, x, t)}{(1-x)\{r(\theta, x, t)\}^2}\, \mathrm{d}x\, \mathrm{d}s \\ &= \int_0^t \int_s^1 \left[\frac{\{1 - \theta(1-x)\} - 2\,(1 - s/x)}{\{r(\theta, x, t)\}^3}\right] \mathrm{d}x\, \mathrm{d}s. \end{aligned}$$

Now for arbitrary $\theta_1, \theta_2 \in (-1 + \delta, 1 - \delta)$ for fixed $0 < \delta < 1$, one finds

$$
\begin{aligned}
&\left|\dot K(\theta_2,t)-\dot K(\theta_1,t)\right| \\
&\quad = \left|\int_0^t\int_s^1 \left\{2\left(1-\frac{s}{x}\right)-1\right\}\left[\frac{1}{\{r(\theta_2,x,s)\}^3}-\frac{1}{\{r(\theta_1,x,s)\}^3}\right]\mathrm{d}x\,\mathrm{d}s\right. \\
&\qquad\left. +\int_0^t\int_s^1 (1-x)\left[\frac{\theta_2}{\{r(\theta_2,x,s)\}^3}-\frac{\theta_1}{\{r(\theta_1,x,s)\}^3}\right]\mathrm{d}x\,\mathrm{d}s\right| \\
&\quad \le 3\int_0^1\int_s^1 \left|\frac{1}{\{r(\theta_2,x,s)\}^3}-\frac{1}{\{r(\theta_1,x,s)\}^3}\right|\mathrm{d}x\,\mathrm{d}s \\
&\qquad +\int_0^1\int_s^1 \left|\frac{\theta_2}{\{r(\theta_2,x,s)\}^3}-\frac{\theta_1}{\{r(\theta_1,x,s)\}^3}\right|\mathrm{d}x\,\mathrm{d}s \\
&\quad \le 4\int_0^1\int_s^1 \left|\frac{1}{\{r(\theta_2,x,s)\}^3}-\frac{1}{\{r(\theta_1,x,s)\}^3}\right|\mathrm{d}x\,\mathrm{d}s \\
&\qquad +|\theta_2-\theta_1|\int_0^1\int_s^1 \frac{1}{\{r(\theta_2,x,s)\}^3}\,\mathrm{d}x\,\mathrm{d}s.
\end{aligned}
$$

Since $r(\theta,x,s)\ge 1-|\theta|>\delta$, the second summand is bounded above by $|\theta_2-\theta_1|/\delta^3$. To handle the first summand, note that $r(\theta,x,s)\le\sqrt{8}$ and $|\dot r(\theta,x,s)|\le 4/\delta$. If follows that

$$
\begin{aligned}
\left|\{r(\theta_2,x,s)\}^3-\{r(\theta_2,x,s)\}^3\right| &= 3\left|\int_{\theta_1}^{\theta_2}\{r(\theta,x,s)\}^2\dot r(\theta,x,s)d\theta\right| \\
&\le \frac{96}{\delta}\,|\theta_2-\theta_1|.
\end{aligned}
$$

Therefore,

$$
\left|\frac{1}{\{r(\theta_2,x,s)\}^3}-\frac{1}{\{r(\theta_1,x,s)\}^3}\right|\le\frac{96}{\delta^7}\,|\theta_2-\theta_1|,
$$

so the first summand is bounded by $384\,|\theta_2-\theta_1|/\delta^7$. Hence hypothesis IV is satisfied for all $\theta\in\mathcal{O}=(0,1)$.

A.3 Asymptotic behavior of the parametric bootstrap method of Wang & Wells (2000)

Without loss of generality, one may assume that $\theta = \tau$. To show that the suggested methodology is incorrect, even in the censored case, let $U_1, \ldots, U_n$ be independent uniformly distributed random variables in $(0,1)$ and for a given value τ_n, set $V^*_{j,n} = K^{-1}_{\tau_n}(U_j)$. Then $V^*_{j,n}$ has distribution $K(\tau_n, \cdot)$.

For simplicity, set $\dot{K}(u) = \partial K(\tau, u)/\partial \tau$, $K(\cdot) = K(\tau, \cdot)$, and $k(\cdot) = k(\tau, \cdot)$. Then, one can show that

$$\dot{Q}(u) = \frac{\partial}{\partial \tau} K^{-1}_{\tau}(u) = -\frac{\dot{K}\left\{K^{-1}(u)\right\}}{k\left\{K^{-1}(u)\right\}},$$

and under appropriate regularity conditions on K, one gets

$$\begin{aligned} E\left\{\dot{Q}(U_j)\right\} &= -\int_0^1 \frac{\dot{K}\left\{K^{-1}(u)\right\}}{k\left\{K^{-1}(u)\right\}}\,\mathrm{d}u = -\int_0^1 \dot{K}(t)\,\mathrm{d}t \\ &= \frac{\partial}{\partial \tau}\int_0^1 \left\{1 - K_\tau(t)\right\}\,\mathrm{d}t = \frac{\partial}{\partial \tau}\left(\frac{\tau+1}{4}\right) = \frac{1}{4}. \end{aligned}$$

Next,

$$K^*_n(t) = \frac{1}{n}\sum_{j=1}^n \mathbf{1}\left\{U_j \le K(\tau_n, t)\right\} = \frac{1}{\sqrt{n}}\,\beta^*_n \circ K(\tau_n, t) + K(\tau_n, t),$$

where

$$\beta^*_n(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^n \left\{\mathbf{1}(U_j \le t) - t\right\}$$

converges in law to a Brownian bridge β^*, independent of $\mathbb{K}_\theta = \beta \circ K - \mu(t, \mathbb{H})$. Hence,

$$\begin{aligned} \sqrt{n}\left\{K^*_n(t) - K\left(\tau^*_n, t\right)\right\} &= \beta^*_n \circ K(\tau_n, t) + \sqrt{n}\left\{K(\tau_n, t) - K\left(\tau^*_n, t\right)\right\} \\ &= \beta^*_n \circ K(\tau_n, t) - \sqrt{n}\left(\tau^*_n - \tau_n\right)\dot{K}(t) + o_P(1). \end{aligned}$$

Moreover, denoting by E_U the expectation with respect to U_j, and setting

$$\bar{\tau}_n = -1 + \frac{4}{n} \sum_{j=1}^{n} K^{-1}(U_j),$$

one obtains

$$\begin{aligned}
\sqrt{n}\,(\tau_n^* - \tau_n) &= \sqrt{n}\left[\frac{4}{n}\sum_{j=1}^{n} K_{\tau_n}^{-1}(U_j) - 4E_U\left\{K_{\tau_n}^{-1}(U_j)\right\}\right] \\
&= \frac{4}{\sqrt{n}}\sum_{j=1}^{n}\left\{K_{\tau_n}^{-1}(U_j) - K^{-1}(U_j)\right\} + \sqrt{n}\,(\bar{\tau}_n - \tau) \\
&\quad -4\sqrt{n}\,E_U\left\{K_{\tau_n}^{-1}(U_1) - K^{-1}(U_1)\right\} \\
&= 4\sqrt{n}\,(\tau_n - \tau)\left\{\frac{1}{n}\sum_{j=1}^{n}\dot{Q}(U_j)\right\} + \sqrt{n}(\bar{\tau}_n - \tau) \\
&\quad -4\sqrt{n}\,(\tau_n - \tau)\,E_U\left\{\dot{Q}(U_1)\right\} + o_P(1) \\
&= \sqrt{n}(\bar{\tau}_n - \tau) + o_P(1).
\end{aligned}$$

It follows that $\sqrt{n}\left\{K_n^*(t) - K\left(\tau_n^*, t\right)\right\}$ converges in distribution to $\beta^* \circ K(t) - Z^*\dot{K}(t)$, where

$$Z^* = -4\int_0^1 \beta^* \circ K(t)\,\mathrm{d}t$$

is the limit in distribution of $\sqrt{n}\,(\tau_n^* - \tau_n)$, since

$$\sqrt{n}\,(\bar{\tau}_n - \tau) = -4\int_0^1 \beta_n^* \circ K(t)\,\mathrm{d}t.$$

Note also that Z^* and the limit Θ of $\sqrt{n}(\tau_n - \tau)$ are independent random variables, but that their distributions are generally different. In fact, $\Theta = 2Z - 4\mathcal{X}_1 - 4\mathcal{X}_2$, where

$$Z = -4\int_0^1 \beta \circ K(t)\,\mathrm{d}t$$

is an independent copy of Z, but

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left\{F_i(X_{ij})-\frac{1}{2}\right\}\rightsquigarrow \mathcal{X}_i, \quad i=1,2.$$

Thus, while the limiting process $I\!\!K^* = \beta^* \circ K - Z^*\dot{K}$ is independent of the limit $I\!\!K = I\!\!K_\tau - \Theta\dot{K}$ appearing in Proposition 2.1, their distributions are clearly not identical, as follows from Barbe et al. (1996). Therefore, the procedure cannot be used to estimate any functional of $I\!\!K_n(\cdot) - K(\tau_n, \cdot)$; in particular it cannot be used to estimate the variance of $S_{\xi n}$ or any P-value.

A.4 Assumptions for the parametric bootstrap based on C_{θ_n}

Here are the conditions under which Genest & Rémillard (2005) prove that the parametric bootstrap procedure described in section 4 is valid:

(R1) For any $\theta \in \mathcal{O}$, the densities c_θ of C_θ exists, are strictly positive on $(0,1)^d$, and are twice differentiable with respect to θ. Moreover, for any $\theta_0 \in \mathcal{O}$,

- $\theta \mapsto \dot{c}_\theta(u)/c_\theta(u)$ and $\theta \mapsto \ddot{c}_\theta(x)/c_\theta(x)$ are continuous at θ_0, for almost every $u \in (0,1)^d$;
- there is a neighborhood $\mathcal{N} = \mathcal{N}(\theta_0)$ of θ_0 such that for all $u \in (0,1)^d$,

$$\sup_{\theta\in\mathcal{N}}\left\|\frac{\dot{c}_\theta(u)}{c_\theta(u)}\right\| \le h_1(u), \qquad \sup_{\theta\in\mathcal{N}}\left\|\frac{\ddot{c}_\theta(u)}{c_\theta(u)}\right\| \le h_2(u),$$

where h_1^2 and h_2 are integrable with respect to C_{θ_0}.

(R2) For any fixed $\theta_0 \in \mathcal{O}$, $\theta \mapsto \dot{k}_\theta(t)$ is continuous at θ_0, for almost all $t \in (0,1)$, and there is a neighborhood $\mathcal{N}$ of θ_0 such that $\sup_{\theta \in \mathcal{N}} \|\dot{k}_\theta(t)\| \leq h_3(t)$, with h_3 integrable over $(0,1)$.

(R3) For any fixed $\theta_0 \in \mathcal{O}$, there exists a square integrable function J_{θ_0}, with respect to C_{θ_0}, such that

$$\theta_n = \frac{1}{n} \sum_{i=1}^{n} J_{\theta_0} \left\{ F_1(X_{1i}), \ldots, F_d(X_{di}) \right\} + o_P \left(1/\sqrt{n} \right),$$

where

$$\int_{(0,1)^d} c_{\theta_0}(u_1, \ldots, u_d) J_{\theta_0} \left(u_1, \ldots, u_d \right) \, \mathrm{d}u_1 \cdots \mathrm{d}u_d = \theta_0,$$

and

$$\int_{(0,1)^d} \dot{c}_{\theta_0}(u_1, \ldots, u_d) J_{\theta_0} \left(u_1, \ldots, u_d \right) \, \mathrm{d}u_1 \cdots \mathrm{d}u_d = I,$$

where I is the $d \times d$ identity matrix.

For example, if the pseudo maximum likelihood of Genest et al. (1995) exists, then it satisfies the regularity assumption R3. If the copula family is indexed by Kendall's tau, then assumption R3 is satisfied. In the latter case, classical nonparametric dependence measures such as Spearman's rho or van der Waerden's coefficient also satisfy R3.

ANNEXE B

B.1 Simulation de copules archimédiennes multivariées

Ici, un algorithme général est proposé pour générer des observations d'une copule archimédienne à $d > 2$ dimensions. En particulier, des formules explicites pour simuler des observations de la copule de Clayton sont obtenues.

Supposons dans ce qui suit que la loi du vecteur $(X_1, \ldots, X_d)$ à $d \geq 2$ dimensions est une copule archimédienne C de générateur ϕ. La distribution jointe conditionnelle de $X_1, \ldots, X_{d-1}$ étant donné que $X_d = w$ est donnée par

$$\begin{aligned} H_w\left(x_1, \ldots, x_{d-1}\right) &= \mathrm{P}\left(X_1 \leq x_1, \ldots, X_{d-1} \leq x_{d-1} \mid X_d = w\right) \\ &= \frac{\partial}{\partial x_d} C(x_1, \ldots, x_d) \\ &= \frac{\phi'(w)}{\phi'\left[\phi^{-1}\left\{\phi(x_1) + \cdots + \phi(x_{d-1}) + \phi(w)\right\}\right]}. \end{aligned}$$

Puisque les distributions marginales univariées de H_w sont

$$F_w(t) = \frac{\phi'(w)}{\phi'\left[\phi^{-1}\left\{\phi(t) + \phi(w)\right\}\right]},$$

une application de la formule (1.2) permet de déduire l'unique copule associée à cette fonction de répartition à $d-1$ variables, à savoir

$$C_w\left(u_1,\ldots,u_{d-1}\right) \;=\; \frac{\phi'(w)}{\phi'\left\{\displaystyle\sum_{i=1}^{d-1}\phi(u_i)+(d-2)\phi(w)\right\}}.$$

La proposition suivante dit que cette copule est archimédienne et le générateur correspondant est donné.

Proposition B. *La copule C_w est archimédienne et son générateur est*

$$\psi_w(t)=\phi\left\{(\phi')^{-1}\left(\frac{\phi'(w)}{t}\right)\right\}-\phi(w), \tag{7.11}$$

où $(\phi')^{-1}$ est l'inverse généralisé de la dérivée première de ϕ.

Exemple B.1. *Pour le modèle paramétrique de Clayton, on calcule que*

$$\psi_w(t)=\frac{w^{-\theta}\left(t^{-\theta/(\theta+1)}-1\right)}{\theta}=k_{\alpha,w}\left(\frac{t^{-\alpha}-1}{\alpha}\right),$$

où

$$\alpha=\frac{\theta}{\theta+1} \quad \text{et} \quad k_{\alpha,w}=(1-\alpha)w^{-\alpha/(1-\alpha)}.$$

Puisque des générateurs identiques, à une constante près, mènent à la même copule archimédienne, la loi de $(X_1,\ldots,X_{d-1})$ étant donné que $X_d=w$ appartient aussi à la famille de Clayton, mais le paramètre de dépendance est réduit à $\alpha=\theta(\theta+1)^{-1}$.

On peut alors dire que la famille de Clayton est invariante par conditionnement. Cette propriété intéressante sera exploitée plus loin afin de proposer un algorithme de simulation de cette loi.

Exemple B.2. *Pour la famille de Frank, quelques calculs montre que le générateur conditionnel vaut*

$$\psi_w(t) = \log\left(\frac{e^{-\theta w}}{t} + 1 - e^{-\theta w}\right).$$

Par conséquent, on conclut que la structure de dépendance de $(X_1, \ldots, X_{d-1})$ *étant donné* $X_d = w$ *appartient à la famille de Ali–Mikhail–Haq de paramètre* $\alpha = 1 - e^{-\theta w}$.

La Proposition B permet d'élaborer un algorithme de simulation pour une copule archimédienne C de générateur ϕ. À cet effet, définissons $\psi^{(0)} \equiv \phi$ et posons que $\psi^{(i)} = \psi_{u_d,\ldots,u_{d-i+1}}$ est le générateur obtenu en conditionnant par rapport à $X_{d-i+1} = u_{d-i+1}$ d'une copule archimédienne dont le générateur est $\psi^{(i-1)}$. On peut montrer par itération que le vecteur $(X_1, \ldots, X_d)$ tel que

$$X_i = \psi^{(0)} \circ \psi^{(d-i)}(U_i), \quad 1 \le i \le d, \tag{7.12}$$

où $U_1, \ldots, U_d$ sont i.i.d. de loi $\mathcal{U}(0,1)$, est de loi C. Dans le cas $d = 2$, on retrouve l'algorithme de simulation

$$\begin{aligned} X_2 &= U_2 \\ X_1 &= \phi^{-1}\left[\phi\left\{(\phi')^{-1}\left(\frac{\phi'(U_1)}{U_2}\right)\right\} - \phi(U_1)\right] \end{aligned}$$

qui est souvent utilisé pour générer des observations bivariées d'une copule archimédienne.

Pour la copule de Clayton, des formules explicites pour (7.12) apparaissent. En effet, on a alors que $\psi^{(0)}(t) = (t^{-\theta} - 1)/\theta$ et on peut montrer, en se basant sur l'exemple B.1, que

$$\psi^{(d-i)}(t) = \frac{1}{\theta}\left(\prod_{j=0}^{d-i-1} u_{d-j}^{-\theta/(j\theta+1)}\right)\left(t^{-\theta/((d-i)\theta+1)} - 1\right).$$

Il découle alors de (7.12) que l'algorithme

$$\begin{aligned} X_d &= U_d \\ X_i &= \left\{\left(\prod_{j=0}^{d-i-1} U_{d-j}^{-\theta/(j\theta+1)}\right)\left(U_i^{-\theta/((d-i)\theta+1)} - 1\right) + 1\right\}^{-1/\theta}, \quad 1 \le i \le d, \end{aligned}$$

où $U_1, \ldots, U_d$ sont i.i.d. $\mathcal{U}(0,1)$, permet de générer un échantillon $(X_1, \ldots, X_d)$ dont la loi jointe est

$$C(x_1, \ldots, x_d) = \left(x_1^{-\theta} + \cdots + x_d^{-\theta}\right)^{-1/\theta}, \quad \theta > 0.$$

B.2 Démonstration de la Proposition 4.3

Du Théorème 3.1 de Klaassen & Wellner (1997), on a la représentation

$$\sqrt{n}\left(r_{jk}^{VdW} - R_{jk}\right) = \frac{1}{\sqrt{n}}\left\{\sum_{i=1}^{n} Z_{ij}Z_{ik} - \frac{R_{jk}}{2}\left(Z_{ij}^2 + Z_{ik}^2\right)\right\} + o_p(1),$$

où

$$(Z_{ij}, Z_{ik})_{i=1}^{n} = \left(N^{-1}\left\{F_j(X_{ij})\right\}, N^{-1}\left\{F_k(X_{ik})\right\}\right)_{i=1}^{n}$$

sont des vecteurs i.i.d. d'une loi normale standard bivariée de corrélation R_{jk}. Ainsi, du Théorème central limite, $W_n = \sqrt{n}(R_n - R)$ converge vers une matrice gaussienne W symétrique, de moyenne nulle, et telle que les éléments de la diagonale principale sont 0. Maintenant, si (Z_j, Z_k, Z_ℓ, Z_q)

sont distribués selon une loi $N_4(\mathbf{0}, R_{jk\ell q})$, avec

$$R_{jk\ell q} = \begin{pmatrix} 1 & R_{jk} & R_{j\ell} & R_{jq} \\ R_{kj} & 1 & R_{k\ell} & R_{kq} \\ R_{\ell j} & R_{\ell k} & 1 & R_{\ell q} \\ R_{qj} & R_{qk} & R_{q\ell} & 1 \end{pmatrix},$$

des calculs élémentaires montrent que

$$\mathrm{E}(Z_j Z_k Z_\ell Z_q) = R_{jk}R_{\ell q} + R_{j\ell}R_{kq} + R_{jq}R_{\ell k}. \tag{7.13}$$

On a donc que

$$\begin{aligned} \mathrm{cov}(W_{jk}, W_{\ell q}) &= \lim_{n\to\infty} \mathrm{cov}\left\{\sqrt{n}\left(r_{jk}^{VdW} - R_{jk}\right), \sqrt{n}\left(r_{\ell q}^{VdW} - R_{\ell q}\right)\right\} \\ &= \mathrm{cov}\left\{Z_j Z_k - \frac{R_{jk}}{2}\,\left(Z_j^2 + Z_k^2\right), Z_\ell Z_q - \frac{R_{\ell q}}{2}\,\left(Z_\ell^2 + Z_q^2\right)\right\}. \end{aligned}$$

Puisque ces deux variables sont chacune de moyenne nulle, le côté droit de l'équation se réduit à

$$\begin{aligned} &\mathrm{E}\left[\left\{Z_j Z_k - \frac{R_{jk}}{2}\,\left(Z_j^2 + Z_k^2\right)\right\}\left\{Z_\ell Z_q - \frac{R_{\ell q}}{2}\,\left(Z_\ell^2 + Z_q^2\right)\right\}\right] \\ &= \mathrm{E}(Z_j Z_k Z_\ell Z_q) - \frac{R_{\ell q}}{2}\,\mathrm{E}\left(Z_j Z_k Z_\ell^2\right) - \frac{R_{\ell q}}{2}\,\mathrm{E}\left(Z_j Z_k Z_q^2\right) - \frac{R_{jk}}{2}\,\mathrm{E}\left(Z_\ell Z_q Z_j^2\right) \\ &\quad - \frac{R_{jk}}{2}\,\mathrm{E}\left(Z_\ell Z_q Z_k^2\right) + \frac{R_{jk}R_{\ell q}}{4}\,\left\{\mathrm{E}\left(Z_j^2 Z_\ell^2\right) + \mathrm{E}\left(Z_j^2 Z_q^2\right) + \mathrm{E}\left(Z_k^2 Z_\ell^2\right) + \mathrm{E}\left(Z_k^2 Z_q^2\right)\right\} \\ &= R_{j\ell}R_{kq} + R_{jq}R_{k\ell} + \frac{1}{2}\,R_{jk}R_{\ell q}\left(R_{j\ell}^2 + R_{jq}^2 + R_{k\ell}^2 + R_{kq}^2\right) \\ &\quad - R_{jk}R_{j\ell}R_{jq} - R_{jk}R_{k\ell}R_{kq} - R_{j\ell}R_{k\ell}R_{\ell q} - R_{jq}R_{kq}R_{\ell q}, \end{aligned}$$

où la dernière égalité s'obtient par des manipulations algébriques simples et un usage répétitif de (7.13). ◊

ANNEXE C

C.1 Proof of Proposition 5.2

First, one needs to show that the expression for σ_J^2 is correct. To this end, set

$$A = \int_{(0,1)^4} \gamma(u,u')\gamma(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v') = A_1 + A_2 + A_3 + A_4,$$

where

$$\begin{aligned}
A_1 &= \int_{0<u\leq u'<1,\ 0<v\leq v'<1} \gamma(u,u')\gamma(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v'),\\
A_2 &= \int_{0<u\leq u'<1,\ 0<v'<v<1} \gamma(u,u')\gamma(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v'),\\
A_3 &= \int_{0<u'<u<1,\ 0<v\leq v'<1} \gamma(u,u')\gamma(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v'),\\
A_4 &= \int_{0<u'<u<1,\ 0<v'<v<1} \gamma(u,u')\gamma(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v').
\end{aligned}$$

Using Tonelli's Theorem, one may write

$$
\begin{aligned}
A_1 &= \int_{0<u\le u'<1,\ 0<v\le v'<1} u(1-u')v(1-v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v') \\
&= \int_{(0,1)^4}\int_{0<u\le u'<1,\ 0<v\le v'<1} \mathbf{1}(x<u)\mathbf{1}(u'\le y)\mathbf{1}(z<v)\mathbf{1}(v'\le w) \\
&\qquad \times \mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w \\
&= \int_{(0,1)^8} \mathbf{1}(x<u\le u')\mathbf{1}(x<u'<y)\mathbf{1}(z<v\le v')\mathbf{1}(z<v'\le w) \\
&\qquad \times \mathbf{1}(x<y)\mathbf{1}(z<w)\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w \\
&= \int_{(0,1)^8} \mathbf{1}(x<u\le u')\mathbf{1}(x<u'\le y)\mathbf{1}(z<v\le v')\mathbf{1}(z<v'\le w) \\
&\qquad \times \mathbf{1}(x<y)\mathbf{1}(z<w)\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w,
\end{aligned}
$$

where the last equality follows from the absolute continuity of Lebesgue's measure.

Similarly,

$$
\begin{aligned}
A_2 &= \int_{0<u\le u'<1,\ 0<v'<v<1} u(1-u')v'(1-v)\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v') \\
&= \int_{(0,1)^4}\int_{0<u\le u'<1,\ 0<v'<v<1} \mathbf{1}(x<u)\mathbf{1}(u'\le y)\mathbf{1}(z<v')\mathbf{1}(v\le w) \\
&\qquad \times \mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w \\
&= \int_{(0,1)^8} \mathbf{1}(x<u\le u')\mathbf{1}(x<u'\le y)\mathbf{1}(v'<v\le w)\mathbf{1}(z<v'<w) \\
&\qquad \times \mathbf{1}(x<y)\mathbf{1}(z<w)\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w \\
&= \int_{(0,1)^8} \mathbf{1}(x<u\le u')\mathbf{1}(x<u'\le y)\mathbf{1}(v'<v\le w)\mathbf{1}(z<v'\le w) \\
&\qquad \times \mathbf{1}(x<y)\mathbf{1}(z<w)\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w.
\end{aligned}
$$

Hence

$$\begin{aligned} A_1 + A_2 &= \int_{(0,1)^8} \mathbf{1}(x < u \leq u')\mathbf{1}(x < u' \leq y)\mathbf{1}(z < v \leq w)\mathbf{1}(z < v' \leq w) \\ &\qquad \times \mathbf{1}(x < y)\mathbf{1}(z < w)\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w. \end{aligned}$$

Using the same technique, one also gets

$$\begin{aligned} A_3 + A_4 &= \int_{(0,1)^8} \mathbf{1}(u' < u \leq y)\mathbf{1}(x < u' \leq y)\mathbf{1}(z < v \leq w)\mathbf{1}(z < v' \leq w) \\ &\qquad \times \mathbf{1}(x < y)\mathbf{1}(z < w)\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w, \end{aligned}$$

from which one may conclude that

$$\begin{aligned} A &= \int_{(0,1)^8} \mathbf{1}(x < u \leq y)\mathbf{1}(x < u' \leq y)\mathbf{1}(z < v \leq w)\mathbf{1}(z < v' \leq w) \\ &\qquad \times \mathbf{1}(x < y)\mathbf{1}(z < w)\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w \\ &= \int_{(0,1)^4} \{J(x,z) + J(y,w) - J(y,z) - J(x,w)\}^2\, \mathbf{1}(x < y)\mathbf{1}(z < w) dxdydzdw \\ &= \frac{1}{4}\int_{(0,1)^4} \{J(x,z) + J(y,w) - J(y,z) - J(x,w)\}^2\, \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w \\ &= \frac{1}{4}\int_{(0,1)^4} \left\{\tilde{J}(x,z) + \tilde{J}(y,w) - \tilde{J}(y,z) - \tilde{J}(x,w)\right\}^2 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}w \\ &= \int_{(0,1)^2} \left\{\tilde{J}(u,v)\right\}^2 \mathrm{d}u\,\mathrm{d}v = \sigma_J^2. \end{aligned}$$

Next, observe that under P_n, it follows by construction that

$$\mathrm{E}\left\{\tilde{\mathbb{C}}_n(u,v)\right\} = 0$$

for any $(u,v) \in [0,1]^2$. Furthermore, for any $(u,v,u',v') \in [0,1]^4$ and $n \geq 2$, one has

$$\mathrm{E}\left\{\tilde{\mathbb{C}}_n(u,v)\tilde{\mathbb{C}}_n(u',v')\right\} = \frac{n}{n-1}\,\gamma_n(u,u')\gamma_n(v,v') \leq \frac{9}{2}\,\gamma(u,u')\gamma(v,v'),$$

where

$$\gamma_n(u,v) = \gamma\left\{C_n(u,1), C_n(v,1)\right\} \le \frac{n+1}{n}\,\gamma(u,v)$$

for arbitrary $(u,v) \in [0,1]^2$.

For any $A \subset (0,1)^2$, let

$$R_{A,n} = \int_A \tilde{\mathbb{C}}_n(u,v)\,\mathrm{d}J(u,v)$$

and define

$$\sigma^2_{A,J} = \int_{A\times A} \gamma(u,u')\gamma(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v').$$

For arbitrary $n \ge 2$, one can then see that under P_n,

$$\begin{aligned}\mathrm{var}\,(R_{A,n}) &= \frac{n}{n-1}\int_{A\times A}\gamma_n(u,u')\gamma_n(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v')\\ &\le \frac{9}{2}\int_{A\times A}\gamma(u,u')\gamma(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v') = \frac{9}{2}\,\sigma^2_{A,J}.\end{aligned}$$

It follows from the Dominated Convergence Theorem that for any $A \subset (0,1)^2$,

$$\lim_{n\to\infty}\mathrm{var}\,(R_{A,n}) = \int_{A\times A}\gamma(u,u')\gamma(v,v')\,\mathrm{d}J(u,v)\,\mathrm{d}J(u',v') = \sigma^2_{A,J} \le \sigma^2_J.$$

In particular, for any $m \ge 1$, one can find a closed interval $K_m \subset (0,1)^2$ so that $K_m \uparrow (0,1)^2$, $\sigma^2_{K_m^c,\,J} < 1/m$ and $\sigma^2_{K_m,\,J} + 1/m > \sigma^2_J$. Hence, for any $\lambda > 0$ and any $n \ge 2$,

$$P_n\left(|R_{K_m^c,\,n}| > \lambda\right) \le \frac{9}{2m\lambda^2}\,.$$

Since m can be chosen arbitrarily large, it follows from the contiguity of Q_n with respect to P_n that for fixed $\lambda > 0$, $\limsup_{n\to\infty} Q_n\left(|R_{K_m^c,\,n}| > \lambda\right)$ may be made arbitrarily small.

Finally, $S_n^J = R_{K_m,\, n} + R_{K_m^c,\, n}$. Moreover, under Q_n, one has

$$R_{K_m,\, n} \rightsquigarrow \int_{K_m} \mathbb{C}(u,v)\,\mathrm{d}J(u,v) + \delta \int_{K_m} \dot{C}_{\theta_0}(u,v)\,\mathrm{d}J(u,v),$$

which is Gaussian, with mean $\delta\mu_{K_m,J}$ and variance $\sigma^2_{K_m,\, J}$. In the light of Condition (iii), it follows that both $\mu_{K_m,J} \to \mu_J$ and $\sigma^2_{K_m,\, J} \to \sigma^2{}_J$, as $m \to \infty$. This completes the proof of Proposition (5.2). ♢

Remark C.1 *Under additional assumptions, e.g., if*

$$\tilde{J}(u,t)\partial\dot{C}(u,t)\Big/\partial u \quad \textit{and} \quad \tilde{J}(t,v)\partial\dot{C}(t,v)\Big/\partial v$$

both converge boundedly to 0 *as* $t \to 1$*, then one may conclude that*

$$\mu_J = \int_0^1 \int_0^1 \tilde{J}(u,v)\dot{c}(u,v)\,\mathrm{d}u\,\mathrm{d}v,$$

as obtained by Genest & Verret (2005), under different assumptions on J*.*

C.2 Proof of Proposition 5.5

For simplicity, set $x = p_\alpha$. It follows from the Gil-Pelaez representation that

$$\beta_B(\delta,\alpha) - \alpha = \frac{1}{2\pi} \int_{-\infty}^{+\infty} t^{-1}\mathrm{Im}\left\{e^{-itx}\hat{f}(t,\delta) - e^{-itx}\hat{f}(t,0)\right\}\,\mathrm{d}t.$$

From the definition of $\hat{f}$, one has

$$t^{-1}\mathrm{Im}\left\{e^{-itx}\hat{f}(t,\delta)\right\} = t^{-1}\xi(t)\,e^{-2\delta^2 t^2\kappa_1(t)}\,\sin\left\{\kappa_2(t) + \delta^2\kappa_3(t) - tx\right\},$$

and it follows that

$$\left(\delta^2 t\right)^{-1}\mathrm{Im}\left\{e^{-itx}\hat{f}(t,\delta) - e^{-itx}\hat{f}(t,0)\right\}$$

can be decomposed as the sum $A_1(t,\delta)t^2\xi(t) + A_2(t,\delta)\xi(t)$, where

$$A_1(t,\delta) = \left(\delta^2 t^3\right)^{-1} \left\{e^{-2\delta^2 t^2 \kappa_1(t)} - 1\right\} \sin\left\{\kappa_2(t) + \delta^2\kappa_3(t) - tx\right\}.$$

and

$$A_2(t,\delta) = \left(\delta^2 t\right)^{-1} \left[\sin\left\{\kappa_2(t) + \delta^2\kappa_3(t) - tx\right\} - \sin\left\{\kappa_2(t) - tx\right\}\right].$$

Now, both terms are bounded and converge respectively, as $\delta \to 0$, to

$$A_1(t,0) = -2t^{-1}\kappa_1(t)\,\sin\left\{\kappa_2(t) - tx\right\}$$

and

$$A_2(t,0) = t^{-1}\kappa_3(t)\,\cos\left\{\kappa_2(t) - tx\right\}.$$

An application of Lebesgue's Dominated Convergence Theorem thus yields

$$\lim_{\delta\to 0} \delta^{-2} \int_{-\infty}^{\infty} t^{-1}\mathrm{Im}\left\{e^{-itx}\hat{f}(t,\delta) - e^{-itx}\hat{f}(t,0)\right\}\,\mathrm{d}t = \int_{-\infty}^{\infty} \psi(t,x)\,\mathrm{d}t,$$

where

$$\psi(t,x) = \xi(t)\left[t^{-1}\kappa_3(t)\cos\{\kappa_2(t) - tx\} - 2t\kappa_1(t)\sin\{\kappa_2(t) - tx\}\right].$$

It is easy to check that ψ can also be expressed as

$$\psi(x,t) = \sum_{k,\,\ell\,\in\,\mathbb{N}} \lambda_{k\ell} I_{k\ell}^2\ \mathrm{Re}\left\{e^{-itx}\hat{f}(t,0)(1 - 2it\lambda_{k\ell})^{-1}\right\}.$$

Since ξ is integrable, it follows that the characteristic function $\hat{f}(t,0)(1 - 2it\lambda_{k\ell})^{-1}$ is integrable, and hence

$$(2\pi)^{-1}\int_{-\infty}^{\infty} \psi(t,x)\,\mathrm{d}t = \pi^{-1}\int_{0}^{\infty} \psi(t,x)\,\mathrm{d}t = \sum_{k,\,\ell\,\in\,\mathbb{N}} \lambda_{k\ell} I_{k\ell}^2 h_{k\ell}(x),$$

where $h_{k\ell}$ is the density of $\mathbb{B}_0 + \lambda_{k\ell}\,\chi_2^2$, whose summands are taken to be independent. This completes the proof. $\diamondsuit$

ANNEXE D

D.1 Tail behavior of a weighted sum of chi-square variables

The following result justifies combination procedure (4) introduced in Section 6.4.

Proposition D.1 *Suppose that* $X = \sum_{k=1}^{\infty} w_k Z_k^2$, *where the random variables* Z_k *are mutually independent and* $Z_k \sim \mathcal{N}(0,1)$, *with* $w_1 \geq w_2 \geq \cdots$, *and* $\mathrm{E}(X) = \sum_{k=1}^{\infty} w_k < \infty$. *Then*

$$\lim_{x\to\infty} \frac{1}{x} \log \mathrm{P}(X > x) = -\frac{1}{2w_1}\,.$$

Proof. Fix $\alpha < 1/(2w_1)$. Then

$$\begin{aligned}
\log\left\{\mathrm{E}\left(e^{\alpha X}\right)\right\} &= -\frac{1}{2}\sum_{k=1}^{\infty} \log\left(1 - 2\alpha w_k\right) \\
&\leq \frac{1}{1-2\alpha w_1}\sum_{k=1}^{\infty} w_k \\
&= \frac{\mathrm{E}(X)}{1-2\alpha w_1} < \infty.
\end{aligned}$$

Hence, by Markov's inequality,

$$\mathrm{P}(X > x) \leq e^{-\alpha x}\mathrm{E}\left(e^{\alpha X}\right)$$

and so

$$\limsup_{x\to\infty} \frac{1}{x} \log \mathrm{P}(X > x) \le -\alpha.$$

Since α can be made arbitrarily close to $1/(2w_1)$, it follows that

$$\limsup_{x\to\infty} \frac{1}{x} \log \mathrm{P}(X > x) \le -\frac{1}{2w_1}\,.$$

One may now conclude, since in view of large deviation results for Gaussian variables, one also has

$$\liminf_{x\to\infty} \frac{1}{x} \log \mathrm{P}(X > x) \ge \liminf_{x\to\infty} \frac{1}{x} \log \mathrm{P}(w_1 Z_1^2 > x) = -\frac{1}{2w_1}\,. \qquad \diamondsuit$$

D.2 An auxiliary result

The following result is instrumental in establishing Proposition 6.7.

Proposition D.2 *Suppose that (w_k) is a positive sequence with $\sum_{k=1}^{\infty} w_k < \infty$ and (μ_k) is a sequence such that $\sum_{k=1}^{\infty} w_k \mu_k < \infty$. Furthermore, let (Z_k) be a sequence of mutually independent random variables with $Z_k \sim \mathcal{N}(0,1)$. For any $\delta \ge 0$, set*

$$X_\delta = \sum_{k=1}^{\infty} w_k (Z_k + \delta\mu_k)^2.$$

Then

$$\lim_{\delta\to 0} \delta^{-2} \left\{ \mathrm{P}\left(X_\delta > x\right) - \mathrm{P}\left(X_0 > x\right) \right\} = \sum_{k=1}^{\infty} w_k \mu_k^2 \, h_k(x),$$

where h_k is a density whose associated characteristic function

$$\frac{\hat{f}(t,0)}{1 - 2iw_k\, t} = (1 - 2iw_k\, t)^{-1} \prod_{j=1}^{\infty} (1 - 2iw_j\, t)^{-1/2}$$

is that of $X_0 + w_k\, \chi_2^2$, in which the summands are taken to be independent.

Proof. It follows from the definition of X_δ that

$$\hat{f}(t,\delta) = \mathrm{E}\left(e^{itX_\delta}\right) = \prod_{k=1}^{\infty} \hat{\eta}\left(w_k\, t, \delta\mu_k\right) = \xi(t)\, e^{-2\delta^2 t^2 \kappa_1(t)}\, e^{i\kappa_2(t)+i\delta^2\kappa_3(t)},$$

where

$$\begin{aligned} \xi(t) &= \prod_{k=1}^{\infty}(1+4t^2w_k^2)^{-1/4}, \quad \kappa_1(t) = \sum_{k=1}^{\infty} w_k^2\mu_k^2/(1+4t^2w_k^2), \\ \kappa_2(t) &= \frac{1}{2}\sum_{k=1}^{\infty}\arctan(2tw_k), \quad \kappa_3(t) = t\sum_{k=1}^{\infty} w_k\mu_k^2/(1+4t^2w_k^2). \end{aligned}$$

Note that $\xi(t)$ and $t^2\xi(t)$ are integrable, that κ_1 is bounded, that $\kappa_i(t)/t$ is bounded for $i=2,3$, and that as $t \to 0$, $\kappa_2(t)/t \to 1/36$ and

$$\frac{\kappa_3(t)}{t} \to \sum_{k=1}^{\infty} w_k\mu_k^2.$$

Next, from the Gil–Pelaez representation, one has

$$\mathrm{P}(X_\delta > x) - \mathrm{P}(X_0 > x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} t^{-1}\mathrm{Im}\left\{e^{-itx}\hat{f}(t,\delta) - e^{-itx}\hat{f}(t,0)\right\}\, dt.$$

Note that

$$t^{-1}\mathrm{Im}\left\{e^{-itx}\hat{f}(t,\delta)\right\} = t^{-1}\xi(t)\, e^{-2\delta^2t^2\kappa_1(t)}\, \sin\left\{\kappa_2(t)+\delta^2\kappa_3(t) - tx\right\}.$$

As a result,

$$\left(\delta^2 t\right)^{-1}\mathrm{Im}\left\{e^{-itx}\hat{f}(t,\delta) - e^{-itx}\hat{f}(t,0)\right\}$$

can be decomposed as the sum $A_1(t,\delta)t^2\xi(t) + A_2(t,\delta)\xi(t)$, where

$$A_1(t,\delta) = \left(\delta^2t^3\right)^{-1}\left\{e^{-2\delta^2t^2\kappa_1(t)} - 1\right\}\sin\left\{\kappa_2(t)+\delta^2\kappa_3(t)-tx\right\}$$

and

$$A_2(t,\delta) = \left(\delta^2 t\right)^{-1}\left[\sin\left\{\kappa_2(t)+\delta^2\kappa_3(t)-tx\right\} - \sin\left\{\kappa_2(t)-tx\right\}\right].$$

Now, both terms are bounded and converge respectively, as $\delta \to 0$, to

$$A_1(t,0) = -2t^{-1}\kappa_1(t)\,\sin\{\kappa_2(t) - tx\}$$

and

$$A_2(t,0) = t^{-1}\kappa_3(t)\,\cos\{\kappa_2(t) - tx\}\,.$$

An application of Lebesgue's Dominated Convergence Theorem thus yields

$$\lim_{\delta\to 0}\delta^{-2}\int_{-\infty}^{\infty} t^{-1}\mathrm{Im}\left\{e^{-itx}\hat{f}(t,\delta) - e^{-itx}\hat{f}(t,0)\right\}\,\mathrm{d}t = \int_{-\infty}^{\infty}\psi(t,x)\,\mathrm{d}t,$$

where $\psi(t,x) = \xi(t)\,[t^{-1}\kappa_3(t)\cos\{\kappa_2(t) - tx\} - 2t\kappa_1(t)\sin\{\kappa_2(t) - tx\}]$.

It is easy to check that ψ can also be expressed as

$$\psi(t,x) = \sum_{k=1}^{\infty} w_k\mu_k^2\;\mathrm{Re}\left\{e^{-itx}\hat{f}(t,0)(1-2itw_k)^{-1}\right\},$$

where $\mathrm{Re}(z)$ stands for the real part of any complex number z.

Since ξ is integrable, it follows that the characteristic function $\hat{f}(t,0)(1 - 2itw_k)^{-1}$ is integrable, and hence

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\infty}^{\infty}\psi(t,x)\,\mathrm{d}t &= \frac{1}{\pi}\int_0^{\infty}\psi(t,x)\,\mathrm{d}t \\
&= \sum_{k=1}^{\infty} w_k\mu_k^2\left(\frac{1}{\pi}\int_0^{\infty}\mathrm{Re}\left\{e^{-itx}\hat{f}(t,0)(1-2itw_k)^{-1}\right\}\,\mathrm{d}t\right) \\
&= \sum_{k=1}^{\infty} w_k\mu_k^2 h_k(x),
\end{aligned}$$

where h_k is the density of $X_0 + w_k\,\chi_2^2$, whose summands are taken to be independent. This completes the proof. $\diamondsuit$

Bibliography

ABDOUS, B., GENEST, C. & RÉMILLARD, B. (2005). Dependence properties of meta-elliptical distributions. In *Statistical Modeling and Analysis for Complex Data Problems*, P. Duchesne & B. Rémillard, eds. Kluwer, Dordrecht, The Netherlands.

ALI, M. M., MIKHAIL, N. N. & HAQ, M. S. (1978). A class of bivariate distributions including the bivariate logistic. *Journal of Multivariate Analysis* **9**, 405–412.

ARTZNER, P., DELBAEN, F., EBER, J.-M. & HEATH, D. (1999). Coherent measures of risk. *Mathematical Finance* **9**, 203–228.

BANDEEN-ROCHE, K. J. & LIANG, K.-Y. (1996). Modelling failure-time associations in data with multiple levels of clustering. *Biometrika* **83**, 29–39.

BARBE, P., GENEST, C., GHOUDI, K. & RÉMILLARD, B. (1996). On Kendall's process. *Journal of Multivariate Analysis* **58**, 197–229.

BEHNEN, K. (1971). Asymptotic optimality and are of certain rank-order tests under contiguity. *The Annals of Mathematical Statistics* **42**, 325–329.

Behnen, K. (1972). A characterization of certain rank-order tests with bounds for the asymptotic relative efficiency. *The Annals of Mathematical Statistics* **43**, 1839–1851.

Belguise, O. & Lévi, C. (2001–2002). Étude des dépendances entre les branches automobile et incendie à l'aide de la théorie des copulas. *Bulletin français d'actuariat* **5**, 135–174.

Bhuchongkul, S. (1964). A class of nonparametric tests for independence in bivariate populations. *The Annals of Mathematical Statistics* **42**, 1839–1851.

Blest, D. C. (2000). Rank correlation — An alternative measure. *Australian and New Zealand Journal of Statistics* **42**, 101–111.

Blum, J. R., Kiefer, J. & Rosenblatt, M. (1961). Distribution free tests of independence based on the sample distribution function. *The Annals of Mathematical Statistics* **32**, 485–498.

Capéraà, P. & Fougères, A.-L. (2001). Estimation of a bivariate extreme value distribution. *Extremes* **3**, 311–329.

Capéraà, P., Fougères, A.-L. & Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* **84**, 567–577.

Chen, X. & Fan, Y. (2005). Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. *La revue canadienne de statistique* **33**, sous presse.

Cherubini, U., Luciano, E. & Vecchiato, W. (2004). *Copula Methods in Finance*. Wiley, New York.

Ciesielska, L. & Ledwina, T. (1983). On locally most powerful rank tests of independence. *Zastosowania Matematyki* **18**, 61–66.

Clayton, D. G. (1978). A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incident. *Biometrika* **65**, 141–151.

Clayton, D. G. & Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model. *Journal of the Royal Statistical Society, Series B* **148**, 82–117.

Cook, R. D. & Johnson, M. E. (1981). A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society, Series B* **43**, 210–218.

Cook, R. D. & Johnson, M. E. (1986). Generalized Burr–Pareto–logistic distributions with applications to a uranium exploration data set. *Technometrics* **28**, 123–131.

Cossette, H., Denuit, M. & Marceau, E. (2002). Distributional bounds for functions of dependent risks. *Bulletin of the Swiss Association of Actuaries* **1**, 45–65.

Cotterill, D. S. & Csörgő, M. (1982). On the limiting distribution of and critical values for the multivariate cramér–von mises statistic. *The Annals of Statistics* **10**, 233–244.

Cotterill, D. S. & Csörgő, M. (1985). On the limiting distribution of and critical values for the multivariate Hoeffding, Blum, Kiefer, Rosenblatt independence criterion. *Statistics and Decisions* **3**, 1–48.

Csörgő, M. (1979). Strong approximations of the Hoeffding, Blum, Kiefer, Rosenblatt multivariate empirical process. *Journal of Multivariate Analysis* **9**, 84–100.

Dakhli, T. (2004). *Analyse de la dépendance de défaut et évaluation des dérivés de crédit sur portefeuille.* Master's thesis, HEC Montréal, Canada.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés: un test non paramétrique d'indépendance. *Académie royale de Belgique, Bulletin de la classe des sciences* **65**, 274–292.

Deheuvels, P. (1980). Non parametric tests of independence. In *Nonparametric Asymptotic Statistics (Proceedings of the Conference held in Rouen in 1979)*, L. N. in Mathematics No 821, ed. Springer, New York.

Deheuvels, P. (1981a). An asymptotic decomposition for multivariate distribution-free tests of independence. *Journal of Multivariate Analysis* **11**, 102–113.

Deheuvels, P. (1981b). A Kolmogorov–Smirnov type test for independence in multivariate samples. *Revue roumaine de mathématiques pures et appliquées* **26**, 212–226.

Deheuvels, P. (1981c). A non parametric test for independence. *Publications de l'Institut de statistique de l'Université de Paris* **26**, 29–50.

Deheuvels, P. & Martynov, G. V. (1996). Cramér–von Mises-type tests with applications to tests of independence for multivariate extreme-value distributions. *Communications in Statistics, Theory and Methods* **25**, 871–908.

Denuit, M., Genest, C. & Marceau, E. (1999). Stochastic bounds on sums of dependent risks. *Insurance: Mathematics and Economics* **25**, 85–104.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R. & Vyncke, D. (2002). The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics* **31**, 3–33.

Dhaene, J., Wang, S., Young, V. & Goovaerts, M. J. (2000). Comonotonicity and maximal stop-loss premiums. *Bulletin of the Swiss Association of Actuaries* , 99–113.

Drouet-Mari, D. & Kotz, S. (2001). *Correlation and Dependence*. London: Imperial College Press.

Dugué, D. (1975). Sur des tests d'indépendance "indépendants de la loi". *Comptes rendus hebdomadaires des séances de l'Académie des sciences* **281**, 1103–1104.

Embrechts, P., Höing, A. & Juri, A. (2003). Using copulae to bound the Value-at-Risk for functions of dependent risks. *Finance and Stochastics* **7**, 145–167.

Embrechts, P., Höing, A. & Puccetti, A. (2005). Worst VaR scenario. *Insurance: Mathematics and Economics* Sous presse.

Embrechts, P., McNeil, A. J. & Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk Management: Value-at-Risk and Beyond*, M. A. H. Dempster, ed. Cambridge, Cambridge University Press.

Fermanian, J.-D., Radulovic, D. & Wegkamp, M. H. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10**, 847–860.

Feuerverger, A. (1993). A consistent test for bivariate dependence. *International Statistical Review* **61**, 419–433.

Fine, J. P., Jiang, H. & Chappell, R. (2001). On semi-competing risks data. *Biometrika* **88**, 907–919.

Frank, M. J. (1979). On the simultaneous associativity of $f(x, y)$ and $x + y - f(x, y)$. *Aequationes Mathematicae* **19**, 194–226.

Frank, M. J., Nelsen, R. & Schweizer, B. (1987). Best-possible bounds on the distribution of a sum – a problem of Kolmogorov. *Probability Theory and Related Fields* **7**, 199–211.

Frank, M. J. & Schweizer, B. (1979). On the duality of generalized infimal and supremal convolutions. *Rend. Mar. Ser. VI* **12**, 1–23.

Frees, E. W. & Valdez, E. A. (1998). Understanding relationships using copulas. *North American Actuarial Journal* **2**, 1–25.

Gänssler , P. & Stute, W. (1987). In *Seminar on Empirical Processes*, DMV Seminar 9. Basel: Birkhäuser.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika* **74**, 549–555.

Genest, C., Ghoudi, K. & Rivest, L.-P. (1995). A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.

Genest, C., Ghoudi, K. & Rivest, L.-P. (1998). Comment on a paper by Frees and Valdez. *North American Actuarial Journal* **2**, 143–149.

Genest, C. & MacKay, R. J. (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *La revue canadienne de statistique* **14**, 145–159.

Genest, C., Marceau, E. & Mesfioui, M. (2002). Upper stop-loss bounds for sums of possibly dependent risks with given means and variances. *Statistics and Probability Letters* **57**, 33–41.

Genest, C. & Plante, J.-F. (2003). On Blest's measure of rank correlation. *La revue canadienne de statistique* **31**, 35–52.

Genest, C., Quessy, J.-F. & Rémillard, B. (2005a). Local efficiency of a Cramér–von Mises test of independence. *Journal of Multivariate Analysis* **94**, sous presse.

Genest, C., Quessy, J.-F. & Rémillard, B. (2005b). On the joint asymptotic behavior of two rank-based estimators of the association parameter in the gamma frailty model. Soumis.

Genest, C. & Rémillard, B. (2004). Tests of independence and randomness based on the empirical copula process. *Test* **13**, 335–369.

Genest, C. & Rémillard, B. (2005). Parametric bootstrap and pseudo-observations. Mimeo.

Genest, C. & Rivest, L.-P. (1993). Statistical inference procedures for bivariate archimedean copulas. *Journal of the American Statistical Association* **88**, 1034–1043.

Genest, C. & Rivest, L.-P. (2001). On the multivariate probability integral transformation. *Statistics and Probability Letters* **53**, 391–399.

Genest, C. & Verret, F. (2005). Locally most powerful rank tests of independence for copula models. *Journal of Nonparametric Statistics* **17**, sous presse.

Ghoudi, K., Khoudraji, A. & Rivest, L.-P. (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *La revue canadienne de statistique* **26**, 187–197.

Ghoudi, K., Kulperger, R. J. & Rémillard, B. (2001). A nonparametric test of serial independence for time series and residual. *Journal of Multivariate Analysis* **79**, 191–218.

Ghoudi, K. & Rémillard, B. (1998). Empirical processes based on pseudo-observations. In *Asymptotic methods in probability and statistics: a volume in honour of Miklós Csörgő*, B. Szyskowitz, ed. Elsevier, Amsterdam.

Ghoudi, K. & Rémillard, B. (2004). Empirical processes based on pseudo-observations II: The multivariate case. In *Asymptotic methods in stochastics: festschrift for Miklós Csörgő*, L. Horváth & B. Szyskowitz, eds., vol. 44 of *The Fields Institute Communications Series*. American Mathematical Society, Providence, RI.

Gieser, P. W. & Randles, R. H. (1997). A nonparametric test of independence between two vectors. *Journal of the American Statistical Association* **92**, 561–567.

Gil-Pelaez, J. (1951). Note on the inversion theorem. *Biometrika* **38**, 481–482.

Glidden, D. V. (1999). Checking the adequacy of the gamma frailty model for multivariate failure times. *Biometrika* **86**, 381–393.

Gumbel, E. J. (1960). Distribution des valeurs extrêmes en plusieurs dimensions. *Publications de l'Institut de statistique de l'Université de Paris* **9**, 171–173.

Hájek, J. & Sidák, Z. (1967). *Theory of Rank Tests*. Academic Press, New York.

He, W. & Lawless, J. F. (2003). Flexible maximum likelihood methods for bivariate proportional hazards models. *Biometrics* **59**, 837–848.

Hennessy, D. A. & Lapan, H. E. (2002). The use of archimedean copulas to model portfolio allocations. *Mathematical Finance* **12**, 143–154.

Hoeffding, W. (1948). A non-parametric test of independence. *The Annals of Mathematical Statistics* **19**, 546–557.

Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika* **48**, 419–426.

Jing, P. & Zhu, L.-X. (1996). Some Blum–Kiefer– Rosenblatt type tests for the joint independence of variables. *Communications in Statistics, Theory and Methods* **25**, 2127–2139.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.

Jouini, M. N. & Clemen, R. T. (1996). Copula models for aggregating expert opinions. *Management Science* **44**, 444–457.

Kaas, R. & Goovaerts, M. J. (1986). Best bounds for positive distributions with fixed moments. *Insurance: Mathematics and Economics* **5**, 87–92.

Kallenberg, W. C. M. & Ledwina, T. (1999). Data-driven rank tests for independence. *Journal of the American Statistical Association* **94**, 285–301.

Klaassen, C. A. J. & Wellner, J. A. (1997). Efficient estimation in the bivariate normal copula model: normal margins are least favorable. *Bernoulli* **3**, 55–77.

Klugman, S. A. & Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics* **24**, 139–148.

Lauprete, G. J., Samarov, A. M. & Welsch, R. E. (2002). Robust portfolio optimization. *Metrika* **55**, 139–149.

Lee, A. J. (1990). *U-Statistics: Theory and Practice*. Marcel Dekker, New York.

Li, D. X. (2000). On default correlation: a copula function approach. *Journal of Fixed Income* **9**, 43–54.

Littell, R. C. & Folks, J. L. (1973). Asymptotic optimality of Fisher's method of combining independent tests II. *Journal of the American Statistical Association* **68**, 193–194.

Makarov, G. (1981). Estimates for the distribution function of the sum of two random variables when the marginal distributions are fixed. *Th. Probab. Appl.* **26**, 803–806.

Marshall, A. W. & Olkin, I. (1988). Families of multivariate distributions. *Journal of the American Statistical Association* **83**, 834–841.

Nelsen, R. B. (1986). Properties of a one-parameter family of bivariate distributions with specified marginals. *Communications in Statistics, Theory and Methods* **15**, 3277–3285.

Nelsen, R. B. (1999). *An Introduction to Copulas.* Lecture Notes in Statistics No 139. Springer, New York.

Nyblom, J. & Mäkeläinen, T. (1983). Comparisons of tests for the presence of random walk coefficients in a simple linear model. *Journal of the American Statistical Association* **78**, 856–864.

Oakes, D. (1982). A model for association in bivariate survival distributions. *Journal of the Royal Statistical Society, Series B* **44**, 414–422.

Oakes, D. (1986). Semiparametric inference in a model for association in bivariate survival data. *Biometrika* **73**, 353–361.

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association* **84**, 487–493.

Oakes, D. (2001). *Biometrika* centenary: Survival analysis. *Biometrika* **88**, 99–142.

Pickands, J. (1981). Multivariate extreme value distributions. *Bulletin de l'Institut international de statistique* **49**, 859–878.

Quesada-Molina, J. J. (1992). A generalization of an identity of Hoeffding and some application. *Journal of the Italian Statistical Society* **3**, 405–411.

Rüschendorf, L. (1982). Random variables with maximum sums. *Advances in Applied Probability* **14**, 623–632.

Ruymgaart, F. H., Shorack, G. R. & van Zwet, W. R. (1972). Asymptotic normality of nonparametric tests for independence. *The Annals of Mathematical Statistics* **43**, 1122–1135.

Scaillet, O. (2005). A Kolmogorov–Smirnov type test for positive quadrant dependence. *La revue canadienne de statistique* **33**, sous presse.

Shih, J. H. (1998). A goodness-of-fit test for association in a bivariate survival model. *Biometrika* **85**, 189–200.

Shih, J. H. & Louis, T. A. (1996a). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **51**, 1384–1388.

Shih, J. H. & Louis, T. A. (1996b). Tests of independence for bivariate survival data. *Biometrics* **52**, 1440–1449.

Shirahata, S. (1974). Locally most powerful rank tests for independence. *Bulletin of Mathematical Statistics* **16**, 11–20.

Shirahata, S. (1975). Locally most powerful rank tests for independence with censored data. *The Annals of Statistics* **3**, 241–245.

Shorack, G. R. & Wellner, J. A. (1984). *Empirical Processes with Applications to Statistics*. Wiley, New York.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de statistique de l'Université de Paris* **8**, 229–231.

Spitzer, F. L. (1974). Introduction aux processus de Markov à paramètre dans z_ν. In *École d'été de probabilités de Saint-Flour III–1973*. Springer, New York.

Tawn, J. A. (1988). Bivariate extreme value theory: models and estimation. *Biometrika* **75**, 397–415.

van den Goorbergh, R. W. J., Genest, C. & Werker, B. J. M. (2005). Bivariate option pricing using dynamic copula models. *Insurance: Mathematics and Economics* **36**, sous presse.

van der Vaart, A. W. & Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.

WANG, J. (2003). The maximum likelihood estimate of the parameters in the logarithmic normal distribution with grouped data. *Acta Mathematicae Applicatae Sinica* **26**, 737–744.

WANG, W. & WELLS, M. A. (2000). Model selection and semi-parametric inference for bivariate failure-time data. *Journal of the American Statistical Association* **95**, 62–72.

WILLIAMSON, R. & DOWNS, T. (1990). Probabilistic arithmetic. I. numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning* **4**, 89–158.